View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by ScholarWorks@UMass Amherst

University of Massachusetts Amherst
ScholarWorks@UMass Amherst

Masters Theses 1911 - February 2014

January 2008

Self-Incompatibility in African Lycium (Solanaceae)

Natalie M. Feliciano
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/theses

Feliciano, Natalie M., "Self-Incompatibility in African Lycium (Solanaceae)" (2008). *Masters Theses 1911 - February 2014*. 133.
Retrieved from https://scholarworks.umass.edu/theses/133

This thesis is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Masters Theses 1911 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

SELF-INCOMPATIBILITY IN AFRICAN *LYCIUM* (SOLANACEAE)

A Thesis Presented

by

NATALIE M. FELICIANO

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE

May 2008

Plant Biology

SELF-INCOMPATIBILITY IN AFRICAN *LYCIUM* (SOLANACEAE)

A Thesis Presented

by

NATALIE M. FELICIANO

Approved as to style and content by:

_____
Jill S. Miller, Chair

_____
Lynn Adler, Member

_____
Ana Caicedo, Member

_____
Elsbeth Walker, Department Head
Plant Biology Graduate Program

DEDICATION

To my brother, David Feliciano, for all your kind words, support and love. Without you, I truly believe that I would not have survived graduate school.

ACKNOWLEDGMENTS

I would like to give an extremely warm and heartfelt thank you to my thesis advisor, Dr. Jill S. Miller. Thank you for the time and energy that you put into helping me with my work and our research. You are an amazing advisor, role model and friend.

Thank you to the North East Alliance for Graduate Education and Professoriate (NEAGEP) for awarding me a graduate fellowship. Thank you to the Dr. Albert L. Delisle Scholarship Fund for awarding me a research grant that helped fund my research travels to South Africa. Thank you to the Botanical Society of America for awarding me a graduate student research award to fund my studies on isolating the *SLF* self-incompatibility gene in *Lycium*.

Thank you to the following southern African permitting agencies: Western Cape Nature Conservation Board, the Free State Department of Tourism, Environmental and Economic Affairs, the Namibian Ministry of Environment and Tourism, the Northern Cape Nature Conservation Service, and the South African National Parks. Additionally, thank you to P. Nel and A. Le Roux for allowing access to the plants used in the pollination experiments.

Chapter one of my thesis has been published in the May 2008 issue of the journal *Evolution* under the title "A TALE OF TWO CONTINENTS: BAKER'S RULE AND THE MAINTENANCE OF SELF-INCOMPATIBILITY IN *LYCIUM* (SOLANACEAE)."

ABSTRACT

SELF-INCOMPATIBILITY IN AFRICAN *LYCIUM* (SOLANACEAE)

MAY 2008

NATALIE M. FELICIANO, B.S., UNIVERSITY OF ROCHESTER

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Dr. Jill S. Miller

Chapter one of my thesis has been published in the May 2008 issue of the journal *Evolution* under the title "A TALE OF TWO CONTINENTS: BAKER'S RULE AND THE MAINTENANCE OF SELF-INCOMPATIBILITY IN *LYCIUM* (SOLANACEAE)." This chapter was co-authored by Dr. Jill S. Miller and Dr. Rachel Levin.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS iv

ABSTRACT v

LIST OF TABLES ix

LIST OF FIGURES xi

CHAPTER

1. A TALE OF TWO CONTINENTS: BAKER'S RULE AND THE MAINTENANCE OF SELF-INCOMPATIBILITY IN *LYCIUM* (SOLANACEAE) 1

1.1 Introduction 1
1.2 Materials and Methods 5

1.2.1 Study System 5
1.2.2 Controlled Pollinations 5
1.2.3 S-RNase Diversity 7
1.2.4 Selection Analyses of the S-RNase Gene 10

1.3 Results 11

1.3.1 Controlled Pollinations 11
1.3.2 S-RNase Diversity 13
1.3.3 Selection Analyses of the S-RNase Gene 14

1.3 Discussion 15

1.3.1 Controlled Pollinations 15
1.3.2 S-RNase Diversity 17
1.3.3 Trans-generic Evolution 18
1.3.4 Selection Analyses of the S-RNase Gene 20

1.4 Conclusions 22

2. ALLELIC DIVERSITY AND SEQUENCE SIMILARITY AMONG S-RNASE ALLELES FROM NINE SPECIES OF OLD WORLD *LYCIUM* 31

2.1 Introduction 31

2.1.1 Gametophytic Self-Incompatibility 31
2.1.2 Study System 35
2.1.3 Self-incompatibility in Old World *Lycium* 36
2.1.4 Current Study 38

2.2 Methods 41

2.2.1 Isolation of *S-RNase* Alleles from Style Tissue 41
2.2.2 Genomic PCR 43
2.2.3 Genealogy and Phylogeny Construction Using Maximum Likelihood 44

2.2.3.1 *S-RNase* Allele Genealogy 44
2.2.3.2 *Lycium* Phylogeny 46

2.2.4 Selection and Diversity at the *S-RNase* Gene 46

2.3 Results 48

2.3.1 Isolation of *S-RNase* Alleles 48
2.3.2 *S-RNase* Genealogy and *Lycium* Phylogeny 49
2.3.3 Selection and Diversity at the *S-RNase* Gene 49

2.4 Discussion 51

2.4.1 S-RNase Genealogy 51
2.4.2 Selection on and Diversity of the *S-RNase* Gene 53
2.4.3 Identical Partial *S-RNase* Alleles 54

2.5 Conclusions 59

3. ISOLATION OF THE *LYCIUM SLF* GENE 73

3.1 Introduction 73
3.2 Methods 76

3.2.1 Study System 76
3.2.2 *SLF* Isolation 77

3.3 Results 79
3.4 Discussion 80
3.5 Conclusions 81

APPENDIX

Raw Data Associated with Fig 1.2 89

BIBLIOGRAPHY 90

LIST OF TABLES

Table Page

1.1. *S-RNase* genotypes for individual *Lycium* sequenced in the present study..... 24

1.2. Likelihood ratio tests comparing models of neutral evolution (M1a and M7) with corresponding models that incorporate positive selection (M2a and M8) for New World and Old World *Lycium* S-RNases..... 25

2.1. Allele specific primers designed to amplify similar *S-RNase* sequences among African *Lycium* species..... 61

2.2. Allele specific primer combinations used to isolate alleles from genomic DNA from selected *Lycium* individuals..... 61

2.3. *S-RNase* alleles, along with their Gen-bank ID numbers, that are included in an alignment of Old World and New World *Lycium*, Solanaceae, and *Antirrhinum S-RNase* sequences..... 62

2.4. Genotypes of Old World *Lycium* 63

2.5. Likelihood ratio test results, parameter values and numbers of positively selected sites for four models of selection [M1a-nearly neutral, M2a-positive, M7(beta)-nearly neutral, M8(beta&ω>1)-positive] in New World and Old World *Lycium*..... 64

2.6. Alleles found in nine southern African *Lycium* species. 65

2.7. The number of individuals, alleles and putatively shared alleles isolated from nine southern African *Lycium* species..... 65

3.1. Allele-specific and degenerate *SLF* primers..... 83

3.2. Gen-bank accession numbers for *Petunia SLF* and putative *Solanum SLF* alleles used in this study 83

3.3. Primer pairs used in PCR amplifications to isolate the *SLF* gene in *Lycium fremontii*, *L. andersonii* and *Petunia inflata*..... 84

3.4. PCR conditions and reagent concentrations for primer combinations..... 85

3.5. Primer pairs used to amplify the *S-RNase* gene in *Lycium fremontii*, *L. andersonii* and *Petunia inflata* 86

3.6. NCBI BLAST (Basic Local Alignment Search Tool) results for cloned *Lycium fremontii* and *L. andersonii* DNA sequences 86

A.1. Raw data associated with Fig. 1.2 89

LIST OF FIGURES

Figure Page

1.1. Schematic illustrating evolutionary relationships among members of tribe Lycieae (A) and Old World species of *Lycium* (B) from Levin et al. (2007) 26

1.2. Means (±1 SE) for fruit production per flower (A) and seed number per fruit (B) in *Lycium ferocissimum* and *L. pumilum* following either outcross (open bars) or self (closed bars) pollination or in unmanipulated controls (shaded bars) 27

1.3. Relationships among S-RNases isolated from Old World (red circles) and New World (blue circles) *Lycium* species, and four additional genera in Solanaceae..... 28

1.4. The distribution of the number of trans-generic lineages as a proportion of topologies inferred from 100 re-sampled data sets 29

1.5. Site-specific selection at the *S-RNase* locus in New World (A) and Old World (B) *Lycium* 30

2.1. Diagram of the *S-RNase* gene in Solanaceae. 66

2.2. Southern African *Lycium* accessions used in this study 66

2.3. Positive amplification of identical partial alleles from genomic DNA 67

2.4. Maximum likelihood S-RNase genealogy constructed using New and Old World *Lycium* and Solanaceae *S-RNase* alleles. 68

2.5. *Lycium* ML phylogeny using nuclear and chloroplast genes..... 69

2.6. Positive selection on the *S-RNase* gene..... 70

2.7. Species ranges for *L. bosciifolium, L. hantamense* and *L. oxycarpum* in Southern Africa 71

2.8. Maximum likelihood phylogeny for *Lycium* using nuclear GBSSI sequences (adopted from Levin et al. 2007) 72

3.1. Isolation of the *Petunia inflata SLF* gene from cDNA using allele specific primers SLF-1f and SLF-8r..... 88

3.2. Positive amplification of *Lycium andersonii* and *Petunia inflata* cDNA using degenerate primers Sd-fi and Sd-ri..... 88

CHAPTER 1

A TALE OF TWO CONTINENTS: BAKER'S RULE AND THE MAINTENANCE OF SELF-INCOMPATIBILITY IN *LYCIUM* (SOLANACEAE)

1.1 Introduction

Both the presence of physiological self-incompatibility and the deployment of sexual function within and among individuals are of profound importance in governing patterns of mating in plant populations (Barrett 2002). Such features control levels of heterozygosity within individuals and genetic diversity in populations, influence the evolution of species interactions, and help maintain reproductive barriers between species. Historically, the most useful systems for studying reproductive evolution are those in which self-incompatibility or gender expression varies either within taxa (e.g., Case and Barrett 2004; Yeung et al. 2005; Barrett and Case 2006) or among closely related species (e.g., Goodwillie 1999; Miller and Venable 2000, 2002; Obbard et al. 2006).

Genetically controlled self-incompatibility systems have evolved repeatedly among angiosperm species and are well established as mechanisms to avoid self-fertilization and prevent inbreeding depression in plants (de Nettancourt 1977; Matton et al. 1994; Castric and Vekemans 2004). In gametophytic self-incompatibility (GSI), two separate but tightly linked genes at the *S*-locus control the recognition specificity. The first is the *S-RNase* gene, which is expressed in the pistil. The second, more recently described gene, is called *SLF* (Sijacic et al. 2004; Qiao et al. 2004) or *SFB* (Ushijima et al. 2003) and is expressed in the haploid pollen grain. When the haploid *S*-genotype of the pollen grain

matches that of either of the two S-RNases expressed in the pistil of the maternal parent, pollen tube growth is terminated and fertilization fails. Given this genetic control, offspring following successful fertilization are necessarily heterozygous. Strong negative frequency-dependence shelters rare alleles from extinction and creates strong selection for novel alleles; both of these factors, the magnitudes of which depend on population size, result in the maintenance of large numbers of *S*-alleles within self-incompatible populations. Several researchers have used allelic diversity at the *S-RNase* gene to characterize mating systems in natural populations (Richman et al. 1996a; Stone and Pierce 2005; Savage and Miller 2006). These studies report heterozygosity and high allelic diversity at this locus, as expected under GSI.

Baker (1955, 1967) was the first to note that self-compatible species are more likely to be successful island colonizers than obligate outcrossers that require pollen transfer between plants (i.e., self-incompatible species). In support of "Baker's Law" (as coined by Stebbins 1957, but see Baker 1967), a higher frequency of self-compatibility, as opposed to self-incompatibility, has been documented in island floras. For example, Anderson et al. (2001) reported that 80% of cosexual species tested in the Juan Fernandez Islands were self-compatible. Likewise, surveys of New Zealand (Webb and Kelly 1993) and the Galapagos Islands (McMullen 1987, 1990) document a similar trend. More recent empirical work has continued to emphasize the generality of Baker's rule in both plants (Schueller 2004; Busch 2005; Flinn 2006) and animals (Trouve et al. 2005), with various exceptions (Carr et al. 1986; Sun and Ritland 1998; Brennan et al. 2005, 2006). Baker also noted the similarity of invasive species to colonizers, and thus predicted invasive species to be disproportionately self-compatible. In southern Africa, Rambuda

and Johnson (2004) found that of 17 invasive species tested, all were self-compatible, and nearly three-quarters were capable of autonomous self-fertilization. Pannell and Barrett (1998) have also modeled the generality of Baker's rule and shown that his ideas apply to plant metapopulations where both colonization and extinction are frequent.

Lycium (Solanaceae) is a genus of ca. 80 species with a cosmopolitan distribution and centers of diversity in Argentina and Chile, southwestern North America, and southern Africa. Recent phylogenetic studies of *Lycium* (Levin and Miller 2005; Levin et al. 2007) strongly support a South American origin of the genus and the monophyly of all Old World species (Fig. 1.1A). Thus, *Lycium* dispersed to the Old World (Africa and Asia) a single time. Extending Baker's (1955, 1967) logic to the dispersal of *Lycium* worldwide, the expectation is that the establishment of *Lycium* in the Old World was facilitated by the presence of self-compatibility in the original colonists.

Members of *Lycium* vary in sexual function; whereas most species are cosexual and produce hermaphroditic flowers, others have gender dimorphism ranging from separate female and hermaphroditic plants (Miller and Venable 2002, 2003) to complete separation of males and females (Venter 2000, 2007). The evolution of gender dimorphism in *Lycium* is especially interesting because it has occurred on a phylogenetic background of self-incompatibility (Bianchi et al. 2000; Miller and Venable 2000, 2002; Richman 2000; Aguilar and Bernardello 2001; see also Fig. 1.1A). Indeed, molecular sequence data for the *S-RNase* gene indicate the presence of shared ancestral polymorphism across several *Lycium* species and other genera in Solanaceae, demonstrating the long-term maintenance of self-incompatibility (Richman and Kohn 2000; Igic et al. 2003; Savage and Miller 2006). The presence of both gender

dimorphism and self-incompatibility presents a challenge to explanations for the evolution of single-sexed plants by an outcrossing advantage, given that all plants in the population are already outcrossing as a result of gametophytic self-incompatibility.

To address this redundancy, Miller and Venable (2000) proposed that gender dimorphism in *Lycium* evolves following polyploidy, which acts as a trigger for the transition to dimorphism because it disrupts the pre-existing self-incompatibility system. The breakdown of self-incompatibility eventually leads to the establishment of single-sexed mutants due to an outcrossing advantage. Consistent with this pathway, all species of *Lycium* with gender dimorphism (n = 3 species in North America; n = 7 species in Africa) are polyploid, whereas cosexual species are diploid (Miller and Venable 2000; Venter 2000). Remarkably, the association between gender dimorphism and polyploidy is also present among populations of a single species, *Lycium californicum*; in this taxon, cosexual populations are diploid, whereas dimorphic populations are polyploid (Yeung et al. 2005). A central assumption of the Miller and Venable (2000) model is the association of self-incompatibility with both diploidy and cosexuality. Although self-incompatibility is documented in diploid, cosexual American *Lycium* (Fig. 1.1A; Richman 2000; Miller and Venable 2002; Savage and Miller 2006), no studies of mating systems in African species have been conducted to date.

Here, we use controlled pollinations to assess the compatibility status for two diploid, cosexual species in Africa, *Lycium ferocissimum* and *L. pumilum*. In addition, we investigate allelic diversity and patterns of selection at the pistil *S-RNase* locus for all alleles retrieved from 15 individuals of five diploid species of southern African *Lycium*. Using these data we address both Baker's assertion that self-compatibility facilitates

establishment, as well as the model proposed by Miller and Venable (2000), which assumes self-incompatibility in diploid, cosexual *Lycium*.

1.2 Materials and Methods

1.2.1 Study System

Lycium (Solanaceae) is a genus of ca. 80 species of predominantly insect-pollinated, perennial shrubs. Within *Lycium*, Levin and Miller (2005) and Levin et al. (2007) have documented a single dispersal event from the Americas to the Old World (Fig. 1.1A). The sister group of *Lycium* is *Nolana* (Olmstead et al. 1999; Levin and Miller 2005), which is estimated to be ca. 10 MY old (Tago-Nakazawa and Dillon 1999); thus, both the origin of *Lycium* and its dispersal to the Old World are relatively recent.

We conducted controlled pollination experiments in natural populations of two African diploid, cosexual species, *Lycium ferocissimum* and *L. pumilum*; the species phylogeny of Old World *Lycium* (Levin et al. 2007; Fig. 1.1B) indicates that these two species are in different clades. We also collected material from five diploid, cosexual species (including *L. ferocissimum*) for analyses of the stylar *S-RNase* gene. These taxa were collected over a wide geographic range and are present in several clades within Old World *Lycium* (Figs. 1.1B,C).

1.2.2 Controlled Pollinations

To determine if *Lycium ferocissimum* and *L. pumilum* are self-compatible, we compared fruit and seed production of flowers pollinated with either self or outcross pollen in natural populations of these species. Pollinations for *L. ferocissimum* were

carried out in West Coast National Park in the Western Cape province, South Africa (S 33° 7.116', E 18° 3.506') during 5–8 September 2006. We pollinated *L. pumilum* from 21–24 September 2006 in Groenefontein Nature Reserve, Western Cape province, South Africa (S 33° 38.266', E 21° 39.129').

On each plant, unopened buds were covered with fine mesh bags to prevent insect visitation. Over the next several mornings, plants were revisited and open flowers with undehisced anthers were emasculated and pollinated with either self (collected from the same plant) or outcross pollen. Outcross pollen was collected from multiple (>10) donors, and each flower in the outcross treatment was pollinated using a minimum of three flowers from this pool. Both pollination treatments delivered sufficient pollen for full seed set. Following pollinations, flowers were re-covered with mesh bags to prevent insect visitation. We pollinated a total of 174 self and 163 outcross flowers on 16 plants of *L. ferocissimum* and 165 self and 136 outcross flowers on 14 plants of *L. pumilum*. In addition, we marked 179 and 177 newly opened, unmanipulated and uncovered flowers to assess natural levels of fruit and seed production in *L. ferocissimum* and *L. pumilum*, respectively. Both populations were revisited and flowers checked for fruit production. Developing fruit were kept covered with fine mesh bags to prevent birds from removing fruit. A total of 172 fruit were collected on 12 November 2006 for *L. ferocissimum*, and 104 fruit were collected from *L. pumilum* on 10 November 2006. To calculate average seed number in the outcross, self, and control treatments, respectively, seed number was counted for all 115, 14, and 43 fruit in *L. ferocissimum* and all 71, 8, and 25 fruit in *L. pumilum*.

Data for fruit production were analyzed using a generalized linear model with a binomial error distribution (individual flowers either succeeded or failed at producing a fruit) and a logit link function (PROC GENMOD, SAS Institute 1989). Each flower was treated independently, and the model included the effects of plant, pollination treatment, and the interaction of plant by pollination treatment. Seed number was analyzed using a general linear model in SAS (PROC GLM, SAS Institute 1989); analyses included the effects of plant, pollination treatment, and the plant by pollination treatment interaction.

1.2.3 S-RNase Diversity

We collected 10-20 styles from five diploid species of African *Lycium*, including four individuals each of *L. oxycarpum* and *L. cinereum*, three individuals of *L. ferocissimum*, and two individuals each from *L. hirsutum* and *L. bosciifolium* (Fig. 1.1C). Styles from mature buds and first-day flowers were preserved in RNAlater® (Ambion, Inc., Austin, TX) and stored initially at 4°C before being transferred to –20°C.

For each individual, we obtained stylar mRNA using the RNeasy Plant Mini Kit (Qiagen, Inc., Valencia, CA) and synthesized cDNA with the First Strand cDNA Synthesis Kit (EMD Biosciences, Inc., Madison, WI). The initial amplification used degenerate primers PR1 and PR3 (Richman et al. 1995) to amplify a portion of the *S-RNase* gene between conserved regions C2 and C5, previously identified by Ioerger et al. (1991) and following conditions in Savage and Miller (2006). Amplification products were cloned into the pT7Blue vector using the Perfectly Blunt Cloning Kit (EMD Biosciences, Inc., Madison, WI). As we do not explicitly demonstrate whether specific sequence variants encode for different specificities, rather than referring to these variants

as “alleles” at the *S-RNase* locus, we will use the terms “sequence variants” or “putative alleles.”

Individual colonies were amplified using the PCR primers and conditions described above. Colony PCR products were screened for restriction fragment length polymorphisms (RFLPs). Approximately twelve colonies were selected from each individual for RFLP analysis, and all colonies with unique RFLP banding patterns were amplified with vector primers U19 (5’-GTT TTC CCA GTC ACG ACG T-3’) and R20 (5’-CAG CTA TGA CCA TGA TTA CG-3’) and sequenced on an Applied Biosystems Automated 3730 DNA Analyzer by the Biotechnology Resource Center at Cornell University (Ithaca, NY). We sequenced an average of seven colonies for each of the 15 accessions. A total of 24 S-RNases were isolated from Old World *Lycium* (Genbank accession numbers: EU074803–EU074826).

To compare Old World and New World *S-RNase* diversity, we included 24 previously published alleles for North American *Lycium parishii* (Savage and Miller 2006; DQ367853-DQ367876) and 11 alleles of North American *L. andersonii*. Only those *L. andersonii* sequences that spanned regions C2-C5 were included in the alignment (Richman 2000; AF05343-4, AF105347-9, AF105353, AF105355, AF105358-9, AF105362-3). In addition, to increase sampling of S-RNases from New World *Lycium*, we sequenced four individuals of the South American species *Lycium cestroides*. These four individuals had a total of six *S-RNase* putative alleles (Genbank accession numbers: EU074797–EU074802).

Sequences were aligned by eye using Ioerger et al. (1991) as a guide and confirmed by comparing multiple sequenced colonies both within and among individuals. All *Lycium*

S-RNases were included in a multiple alignment that included 55 S-RNases from previously published studies of Solanaceae: *Petunia axillaris* (AF239907–10, AY180048, AY180050), *Petunia integrifolia* (AF301167–8, AF301171–3, AF301176–7, AF301180), *Physalis cinerascens* (AF058930–1, AF058933, AF058935–7, AF058940), *Physalis crassifolia* (L46653, L46656–8, L46663, L46668-9, L46672–3), *Solanum carolinense* (L40539–46), *S. chacoense* (AF176533, L36666, S69589, X56896–7), *Witheringia maculata* (AF102066–7, AF102070–3), *W. solanacea* (AY454099, AY454105, AY454111, AY454113–4, AY454117). The *S-RNase* S2 from *Antirrhinum hispanicum* (X96465) was used as the outgroup.

Average pairwise distances for S-RNases from Old and New World *Lycium* were calculated in PAUP* (Swofford 2002). We constructed gene genealogies of the complete 120 S-RNase data set using both maximum likelihood and Bayesian approaches. Maximum likelihood (ML) model parameters were determined using the Akaike information criterion in Modeltest version 3.7 (Posada and Crandall 1998). The best-fit model (GTR+I+G) was used in an ML analysis in PAUP* using the heuristic search option, TBR branch swapping, MulTrees option in effect, and a single neighbor-joining tree as a starting topology. The Bayesian analysis was run in MrBayes version 3.1.2 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003) and included four simultaneous Markov chain Monte Carlo (MCMC) chains, each starting from a random tree, with a general time reversible substitution model and gamma-distributed rate variation across sites. Two million generations were run, with a tree saved every 100 generations; trees preceding the stabilization of likelihood values were excluded in the construction of a consensus tree in PAUP*.

The number of trans-generic lineages for Old World and New World *Lycium* was determined for both the ML and Bayesian topologies; we define a trans-generic lineage (TGL) as the most recent node including sequences from *Lycium* and another genus. However, because the number of TGLs is sensitive to the number of sequences included, it is possible that a greater number of TGLs will be inferred for New World *Lycium* as a result of a larger sample (41 New World S-RNases versus 24 Old World S-RNases). In addition, estimates of TGLs are also sensitive to uncertainty in the genealogy. To investigate these concerns, we used a re-sampling approach in which we randomly selected 24 (of the 41 available) New World S-RNases. These 24 randomly selected New World sequences, all 24 Old World sequences, and the 55 S-RNase Solanaceae-wide data set were then used to construct a neighbor-joining tree in PAUP*, and the number of TGLs for New World and Old World *Lycium* was determined. This procedure was repeated 100 times.

1.2.4 Selection Analyses of the S-RNase Gene

As positive selection occurs in particular regions of the *S-RNase* gene (Savage and Miller 2006), we investigated patterns of selection using site-specific codon models as implemented in the *codeml* package in PAML version 3.15 (Yang 1997). Specifically, we used *codeml* for estimates of the d_N/d_S rate ratio (ω) for New World and Old World *Lycium* S-RNases. Likelihood ratio tests (LRTs) were used to compare null models that assume nearly-neutral evolution (M1a and M7) with more complex models that incorporate positive selection (M2a and M8; models in Yang et al. 2000 and Wong et al. 2004). The use of both likelihood ratio tests (M1a versus M2a and M7 versus M8) is the recommended procedure for investigating positive selection using PAML (Yang and

Bielawski 2000; Wong et al. 2004), and similarity between analyses can be used to assess robustness. The M1a model assumes two amino acid site classes; one site class fixes ω at one, whereas ω for the second site class is estimated under the constraint that it varies between zero and one (models in Wong et al. 2004). Thus, in the M1a model, sites are either evolving neutrally or nearly-neutrally ($\omega_1 = 1$) or are subject to purifying selection (i.e., under the constraint, $0 < \omega_0 < 1$). The M2a positive selection model has a third site class (ω_2) that is also estimated; ω_2 can exceed one and is the d_N/d_S rate ratio of those sites under positive selection. We also compared models M7 (beta) and M8 (beta&ω>1; models in Yang et al. 2000). Both M7 and M8 assume a β-distribution for $0 \leq \omega \leq 1$, but model M8 includes an additional parameter that incorporates positive selection. Where LRTs indicated that the models incorporating positive selection (M2a and M8) fit the data significantly better than their corresponding null models (M1a and M7) and the d_N/d_S rate ratio associated with positive selection (ω_2 in model M2a and ω_s in model M8) was greater than one, diversifying selection was inferred for sites in the sequences. The empirical Bayesian probabilities (Yang et al. 2005) were then used to determine the number and identity of sites under positive selection.

1.3 Results

1.3.1 Controlled Pollinations

For both *Lycium ferocissimum* and *L. pumilum*, outcross pollination resulted in greater fruit and seed production compared to self pollination. There was a significant main effect of pollination treatment on fruit production for both species (*L. ferocissimum*, $\chi^2 = 110.9$, df = 2, $P < 0.0001$; *L. pumilum*, $\chi^2 = 72.2$, df = 2, $P < 0.0001$). In *L. ferocissimum*,

75% of flowers pollinated with outcross pollen set fruit compared to only 8% following self pollination. Likewise, in *L. pumilum* 60% of flowers formed fruit following outcross pollination, whereas only 5% of self pollinations resulted in fruit production (Fig. 1.2A). There was also a main effect of plant for both species (*L. ferocissimum*, $\chi^2 = 58.3$, df = 15, $P < 0.0001$; *L. pumilum*, $\chi^2 = 41.1$, df = 13, $P < 0.0001$), indicating that some plants had higher fruit production than others, regardless of pollination treatment. Outcross pollination was also over 2.5 times more effective at fruit production compared to unmanipulated control flowers (0.75 vs. 0.27 for *L. ferocissimum* and 0.60 vs. 0.18 for *L. pumilum*; Fig. 1.2A). There was a significant plant by pollination treatment for *L. ferocissimum* ($\chi^2 = 51.6$, df = 30, $P = 0.009$), but not *L. pumilum* ($\chi^2 = 26.9$, df = 26, $P = 0.414$). Comparison of the plant by pollination treatment means for *L. ferocissimum* revealed that the outcross pollination treatment was more successful than the self pollination treatment for all genotypes (the difference between outcross and self pollination treatments ranged from 0.48–1.0). The only exception, and the source of the interaction, was a single individual in which fruit set for both the outcross and self pollination treatments was low (both treatments, 0.09; control fruit set was 0.20 for this plant); removal of this individual and re-analysis resulted in a non-significant interaction.

For seed production in both species, there were significant main effects for both pollination treatment (*L. ferocissimum*, $F_{2,39.2} = 22.3$, $P < 0.0001$; *L. pumilum*, $F_{2,30.2} = 60.9$, $P < 0.0001$) and plant (*L. ferocissimum*, $F_{15,21.2} = 2.3$, $P < 0.039$; *L. pumilum*, $F_{13,17.7} = 4.5$, $P < 0.002$). On average, outcrossed fruits produced roughly five and a half times more seed than selfed fruits for *L. ferocissimum* (43.1 vs. 7.9) and over seven and a half times as many for *L. pumilum* (14.7 vs. 1.9; Fig. 1.2B). Seed set in unmanipulated

control fruits was significantly higher than for self fruits, but lower than outcrossed seed number in both species (Fig. 1.2B).

1.3.2 S-RNase Diversity

Sequences ranged from 363–384 bp in length and 30 unique *S-RNase* sequence variants were identified. Twenty-four putative alleles were recovered among the 15 individuals of Old World *Lycium*, and six putative alleles were designated in the four New World *L. cestroides* individuals (Table 1.1). Consistent with the presence of self-incompatibility, thirteen of fifteen Old World individuals and all of the South American individuals were heterozygous at the *S-RNase* locus.

Old World S-RNases were more similar to each other than were comparisons within New World S-RNases. Specifically, average pairwise amino acid distances were 0.34 and 0.53 for Old World and New World S-RNases, respectively. There was a group of similar S-RNases among the Old World putative alleles (16 sequences with average AA distance of 0.075). Although average divergence was low in this group, all had unique combinations of amino acids (average difference among these sequence variants was 9.4 amino acids). The only exception was the comparison between *OXYC6* and *CINE4*, which differed at only 2 synonymous sites; however, these differences were confirmed using multiple independent clones.

Results of the ML and Bayesian analyses were similar, and the ML genealogy is presented in Fig. 1.3. The 65 *Lycium* S-RNases occur in 11 TGLs; however, New World and Old World sequences were not distributed equally among these lineages. S-RNases isolated from New World species were present in all of the TGLs, whereas those from Old World taxa were present in only four TGLs (Fig. 1.3). Re-sampling using equal

numbers of New and Old World S-RNases reinforced this finding; the number of S-RNase TGLs in the Old World was only half the number found in the New World (Fig. 1.4).

1.3.3 Selection Analyses of the S-RNase Gene

Selection analyses indicated that the model of positive selection (M2a) fit the data significantly better than the model of nearly-neutral (M1a) evolution for both the Old World and New World samples (Old World: LRT = 48.08, $P < 0.0001$; New World: LRT = 15.14, $P < 0.001$; Table 1.2). Likewise, the second set of LRTs showed the same pattern, with the model incorporating positive selection (M8) fitting the data significantly better than the null model (M7) (Old World: LRT = 51.83, $P < 0.0001$; New World: LRT = 16.56, $P < 0.001$).

Not surprisingly, the majority of positively selected sites for both data sets were located in the hypervariable regions of the *S-RNase* gene (Fig. 1.5). However, the number of positively selected sites among Old World S-RNases was 3–4 times higher than the estimate for New World S-RNases (Fig. 1.5). In addition, the d_N/d_S rate ratio for those sites in the positive selection class (ω_2 for the M2a model and ω_s for the M8 model; Table 1.2) ranged from 4.4–4.7 for Old World putative alleles, but from only 1.4–1.9 for the New World data set.

1.3 Discussion

1.3.1 Controlled Pollinations

The two South African species of *Lycium* studied here, *Lycium ferocissimum* and *L. pumilum*, are strongly self-incompatible based on fruit and seed production in controlled crosses (Fig. 1.2). In our pollinations, outcrossing resulted in a 51- (*L. pumilum*) or 93-fold (*L. ferocissimum*) increase in seed production per flower compared to selfing. The present data for African species indicate stronger self-incompatibility than recorded previously for North American species of *Lycium* (Miller and Venable 2000, 2002). Miller and Venable (2002) documented a 14 to 15-fold advantage of outcross, compared to self pollen, for three North American species. It is not clear why the relative success of outcross and self pollen should vary so dramatically among species. Data from two of three North American species suggests higher fruit set following self-pollination (Miller and Venable 2002) than reported here for the South African species. An influencing factor includes the timing of pollination, which is known to affect self fruit set in self-incompatible species (Stone 2004; Travers et al. 2004). Additionally, in the present study we pooled pollen from > 10 individuals to serve as outcross pollen donors, whereas Miller and Venable (2002) in some cases used as few as two outcross donors. As the genotypes of individuals were unknown at the time of pollination in both studies, it is possible that the larger pool of outcross pollen in the present study increased the likelihood of compatible crosses among outcross pollinations. Paschke et al. (2002) experimentally investigated the effects of pollen load diversity in self-incompatible *Cochlearia bavarica* (Brassicaceae) and reported that increasing the number of pollen donors (i.e., pollen diversity) was associated with higher reproductive success.

Although the presence of self-incompatibility in African *Lycium* appears to contradict Baker's rule (1955, 1967), there is ample support for Baker's assertions in the literature. Much of this evidence involves broad scale comparative surveys of island floras, and many such studies document the relative rarity of self-incompatibility compared to self-compatibility on islands (McMullen 1987, 1990; Webb and Kelly 1993; Bernardello et al. 2001; Anderson et al. 2001). Likewise, Rambuda and Johnson (2004) document self-fertility among alien invasive species in South Africa. However, evidence within species or among closely related species provides more powerful tests of Baker's assertions. In *Lycopersicon hirsutum* (=*Solanum habrochaites*), populations are self-incompatible in the center of the species range, but self-compatibility is thought to have evolved independently in peripheral northern and southern populations (Rick et al. 1979; Rick and Chetelat 1991). More recently, Busch (2005) and Ortiz et al. (2006) have found increased selfing rates in peripheral populations, as opposed to central populations, consistent with the hypothesis that self-fertility facilitates colonization and range expansion. Likewise, Schuller (2004) investigated the capacity for self-fertilization in island and mainland populations of self-compatible *Nicotiana glauca*. Her data suggest that enhanced self-fertilization in island (versus mainland) populations is a result of the establishment success of highly self-fertilizing genotypes, as opposed to selection for selfing following colonization. A recent study by Flinn (2006) also documents the association of self-fertilization ability and colonization history; specifically, among three fern species those with frequent colonization had the highest selfing rates.

In contrast, only a handful of studies (including the results presented here) demonstrate clear exceptions to Baker's rule. For example, Brennan et al. (2005, 2006)

have documented the maintenance of sporophytic self-incompatibility in *Senecio squalidus* (Asteraceae) following its establishment and expansion into Britain. Brennan et al. (2002) also suggested that the presence of pseudo-incompatibility in some genotypes might have aided this expansion. Sun and Ritland (1998) estimated high outcrossing rates for introduced populations of *Centaurea solstitialis* (Asteraceae), although the genetic control of incompatibility in this system is unknown. Other groups that have maintained self-incompatibility after colonization include the silversword alliance radiation in Hawaii (Carr et al. 1986) and the colonizing, clonal species *Ipomoea pes-caprae* (Convolvulaceae; Devall and Thien 1992). It is notable that published exceptions are from species with sporophytically controlled self-incompatibility systems as found in Asteraceae and Convolvulaceae. Our data for *Lycium* also suggest the maintenance of self-incompatibility following long-distance dispersal, and to our knowledge this is the only documented exception to Baker's rule in a species with gametophytically controlled self-incompatibility. Although it is not clear if differences between the two incompatibility systems should be expected, some authors have suggested that selection for increased dominance interactions among alleles can allow for greater mate availability under sporophytically controlled self-incompatibilty (Byers and Meagher 1992; Brennan et al. 2003), whereas in gametophytic systems, dominance among incompatibility alleles is unknown.

1.3.2 S-RNase Diversity

We isolated 24 unique *S-RNase* sequences from 15 individuals and five species of African *Lycium*. Of the 19 individuals genotyped in the present study, 17 were heterozygous (Table 1.1). The putative alleles recovered for the African taxa were more

similar (average pairwise amino acid identity 66%) than equivalent comparisons among New World alleles (48% among North American *L. parishii* alleles, Savage and Miller 2006; 44% among South American *L. cestroides*, this study). However, the estimate of nucleotide diversity (π) for Old World alleles (MEGA version 3.1, Kumar et al. 2004) was 0.330 and of similar magnitude to π calculated for other species with gametophytic self-incompatibility (ranges from 0.208–0.463 in Table 1 of Castric and Vekemans 2004).

Four of the five *Lycium ferocissimum* putative alleles are included in a large group of closely related Old World S-RNases (top clade in Fig. 1.3); these four sequences differ from each other at an average of 8.7 (range 7–12) amino acid sites. Although quite similar, the majority of their differences lie within the hypervariable regions, which are known to be associated with specificity determination (Ioerger et al. 1991; Ida et al. 2001). Further, Matton et al. (1999) and Wang et al. (2001) have noted that changes in as few as five amino acid sites are sufficient to generate novel specificities. Our *S-RNase* data for *L. ferocissimum*, coupled with the crossing data demonstrating self-incompatibility in this species (Fig. 1.2), provide strong evidence that similar *S-RNase* sequence variants can have distinct functional phenotypes. That said, we obtained only partial *S-RNase* sequences, and more extensive sequencing could reveal additional differences.

1.3.3 Trans-generic Evolution

Igic et al. (2003) documented that gametophytic self-incompatibility is the ancestral condition in Solanaceae; in fact, GSI is ancestral for nearly three-quarters of eudicot lineages (Igic and Kohn 2001; Steinbachs and Holsinger 2002). Gametophytic self-incompatibility is a complex character requiring the coordinated action of several genes

(McClure 2006) and, as such, it has been argued that this trait is more likely to be lost than gained over evolutionary time (Igic et al. 2006). The regain of self-incompatibility following its loss has never been documented in Solanaceae (Igic et al. 2006; see also Takebayashi and Morrell 2001), despite many empirical studies of the *S-RNase* locus among genera in this family (Ioerger et al. 1990; Richman et al. 1996a,b; Richman and Kohn 2000; Lu 2001; Wang et al. 2001; Igic et al. 2003; Stone and Pierce 2005; Savage and Miller 2006). Following the loss of self-incompatibility, polymorphism at the *S*-locus is expected to decline rapidly, because functionally distinct alleles are rendered selectively neutral (e.g. Igic et al. 2003, 2006; Charlesworth and Vekemans 2005). Empirical studies of *S-RNase* diversity following the transition to self-compatibility support this loss of allelic diversity (Golz et al. 1998; Kondo et al. 2002). If self-incompatibility were to be regained following its loss, then S-RNases in descendant taxa would lack trans-generic polymorphism (Igic et al. 2006).

In contrast, our finding of multiple trans-generic *S-RNase* lineages in Old World *Lycium* is consistent with the presence of self-incompatibility in the colonists. The majority of Solanaceae studied to date possess deep coalescence of S-RNases with many trans-generic lineages (Richman and Kohn 2000; Igic et al. 2003). Consistent with previous studies of self-incompatible *Lycium* species (Richman 2000; Savage and Miller 2006), *Lycium* S-RNases in the present study are distributed among many trans-generic lineages. However, S-RNases from New and Old World species are not distributed evenly. Whereas S-RNases isolated from New World species are present in all of the TGLs (11 TGLs in Fig. 1.3), those from Old World taxa are present in only four TGLs (Fig. 1.3). Further, this result is robust to subtleties in topology, as well as differences in

sampling between geographic regions. Our results show clearly that the number of S-RNase TGLs in the Old World is at most half (4 TGLs) of the number recovered for New World sequences (mean = 10.04, 95% CI 9.8–10.2; Fig. 1.4). Further, Old World S-RNases are nested within S-RNases from New World *Lycium* (Fig. 1.3), a pattern expected if dispersal to and colonization of the Old World occurred relatively recently. Figure 1.3 shows that there are six "trans-geographic" lineages (i.e., well supported lineages including both Old and New World *Lycium* S-RNases), which implies a minimum of three colonists (each with two different New World allelic lineages). However, we note that uncertainty in the genealogy, as well as limited *S-RNase* sampling, will affect this estimate; additional sampling and simulations may allow for more robust estimates of the initial population size of colonists.

This pattern of reduced TGLs in Old World *Lycium* is similar to that observed in *Physalis* (Richman and Kohn 2000) and *Witheringia* (Stone and Pierce 2005) species, which together share a reduced set of trans-generic lineages (see also Fig. 1.3). Richman and colleagues (1996b) have suggested that a population bottleneck occurred in the common ancestor of *Physalis* and *Witheringia*, a result consistent with their close evolutionary relationship (Olmstead et al. 1999). Despite the loss of considerable trans-generic diversity in these taxa, allelic diversity persists and demonstrates the maintenance of gametophytic self-incompatibility following a bottleneck (Richman et al. 1996b).

1.3.4 Selection Analyses of the S-RNase Gene

Positive selection can be difficult to detect in coding genes, given that most sites are expected to be under purifying selection and relatively few sites are expected to undergo positive selection. The site-specific analyses implemented in PAML have proven useful

for inference of positive selection in the *S-RNase* gene (Takebayashi et al. 2003; Savage and Miller 2006; Igic et al. 2007) and other coding regions (Yang and Bielawski 2000). Our selection analyses indicate positive selection at certain sites in the *S-RNase* gene for both Old and New World S-RNases (Table 1.2; Fig. 1.5). Interestingly, comparison of the Old and New World data sets indicates relatively intense positive selection in Old World alleles. Specifically, the proportion of sites falling into the positive selection class was higher for the Old World compared to New World S-RNases (model M2a: p_2 = 0.199 versus 0.102, respectively; Table 1.2). Further, the d_N/d_S rate ratio was higher for Old World compared to New World alleles (model M2a: ω_2 = 4.71 versus 1.88, respectively; model M8: ω = 4.43 versus 1.43; Table 1.2). Thus, it appears that intense positive selection has occurred among Old World S-RNases. Such a result may be indicative of the re-diversification of Old World S-RNases following colonization and establishment.

Given that the average amino acid pairwise distance was only 7.5% among the similar group of Old World S-RNases (top clade in Fig. 1.3), we also used PAML to analyze S-RNases in this clade to examine patterns of selection in these closely related sequences. Our analysis rejects the hypothesis that Old World S-RNases in this large clade are evolving neutrally (M1a versus M2a: LRT = 38.245, $P < 0.0001$; M2a parameters, p_2 = 0.09 and ω_2 = 12.26; M7 versus M8: LRT = 38.280, $P < 0.0001$; M8 parameters, p_1 = 0.089 and ω_s = 12.43). Thus, despite their similarity, these S-RNases have an excess of nonsynonymous substitutions as measured by the d_N/d_S rate ratio. As across the data set for all Old World *Lycium*, analysis of the large group of Old World sequences suggests

positive selection favoring new specificities, consistent with the re-diversification of alleles following a founder event.

1.4 Conclusions

Baker (1955, 1967) was the first to suggest that self-fertility facilitates the successful establishment of species following long-distance dispersal. Although the transition from self-incompatibility to self-compatibility has occurred often in the history of Solanaceae (Igic et al. 2006), our data for *Lycium* demonstrate an exception to Baker's ideas, indicating instead the maintenance of gametophytic self-incompatibility following dispersal from the New World to the Old World. However, the trans-generic diversity of S-RNases in Old World *Lycium* is limited in comparison to New World alleles, a result consistent with a genetic bottleneck coincident with dispersal of *Lycium* to the Old World. These data are some of the first from a natural population to suggest that allelic differences in as few as 7–8 amino acids are sufficient for generation of novel allele specificities.

It is interesting to speculate about factors that may have facilitated the maintenance of self-incompatibility in Old World *Lycium*. Species of *Lycium* are typically shrubs that produce many-seeded yellow to red (sometimes black) fleshy berries, which are generally bird dispersed. Although repeated migration between southwestern North America and Argentina and Chile has been suggested (Levin and Miller 2005), all Old World *Lycium* sampled to date comprise a monophyletic group. In addition, all *Lycium* are long-lived perennials, and Bowers (2005) and Bowers et al. (1995) have recorded considerable longevity (120–211 years) for species of southwestern North American *Lycium*. These factors may well have aided the retention of self-incompatibility in *Lycium*, as a single

multi-seeded fruit could carry several incompatibility alleles; likewise, the presence of longevity increases opportunities for mating among colonists.

Although there have been a number of studies documenting self-incompatibility in American *Lycium* (Richman 2000; Aguilar and Bernardello 2001; Miller and Venable 2002; Savage and Miller 2006), this study is the first to examine breeding systems among Old World *Lycium* species. The finding of self-incompatibility in two diploid, cosexual African *Lycium* species, *L. ferocissimum* and *L. pumilum*, is consistent with a model for the evolution of gender dimorphism proposed by Miller and Venable (2000), which assumes self-incompatibility in cosexual, diploid species. In this model, gender dimorphism evolves following the loss of self-incompatibility in diploid, cosexual species. Polyploidy is implicated as the trigger for the loss of self-incompatibility, and considerable support for the breakdown of incompatibility with polyploidy exists in families with GSI (see references in Miller and Venable 2000, 2002, and Table 1 in Mable 2004). Among American species of *Lycium*, available data are consistent with the evolution of gender dimorphism as proposed by Miller and Venable (2000). Results of the present study suggest that this model may also explain the distribution of cosexual and dimorphic species in African *Lycium*.

Table 1.1. *S-RNase* genotypes for individual *Lycium* sequenced in the present study.

Species	Genotype
Old World	
L. bosciifolium	*BOSC1 / BOSC4*
L. bosciifolium	*BOSC2 / BOSC3*
L. cinereum	*CINE1 / CINE2*
L. cinereum	*CINE1 / CINE5*
L. cinereum	*CINE3 / CINE4*
L. cinereum	*CINE6*
L. ferocissimum	*FERO1 / FERO2*
L. ferocissimum	*FERO2 / FERO5*
L. ferocissimum	*FERO3 / FERO4*
L. hirsutum	*HIRS1 / HIRS2*
L. hirsutum	*HIRS3*
L. oxycarpum	*OXYC1 / OXYC5*
L. oxycarpum	*OXYC2 / OXYC3*
L. oxycarpum	*OXYC2 / OXYC3*
L. oxycarpum	*OXYC4 / OXYC6*
New World	
L. cestroides	*CESTa / CESTb*
L. cestroides	*CESTa / CESTf*
L. cestroides	*CESTb / CESTe*
L. cestroides	*CESTc / CESTd*

Table 1.2. Likelihood ratio tests comparing models of neutral evolution (M1a and M7) with corresponding models that incorporate positive selection (M2a and M8) for New World and Old World *Lycium* S-RNases. For models M1a and M2a, values of p_0, p_1, and p_2 are the proportion of sites inferred to be evolving under purifying selection, neutral evolution, and positive selection, respectively, and ω_0, ω_1, *and* ω_2 are their corresponding d_N/d_S rate ratios. For models M7 and M8, the beta distribution, beta(p,q), describes the distribution of the d_N/d_S rate ratio between zero and one, and p_0 is the proportion of sites within this distribution. In these models, p_1 is the proportion of sites inferred to be under positive selection, and ω_s is the d_N/d_S rate ratio for those sites.

	Model	ℓ	$2\Delta\ell$	**Parameters**
New World *Lycium*	M1a: nearly neutral	–7430.833		$p_0 = 0.449$ ($\omega_0 = 0.202$); $p_1 = 0.551$ ($\omega_1 = 1$)
	M2a: positive selection	–7423.265	15.14^{***}	$p_0 = 0.432$ ($\omega_0 = 0.213$); p_1=0.466 ($\omega_1 = 1$) $p_2 = 0.102$ ($\omega_2 = 1.88$)
	M7: beta	–7410.260		$p = 0.554$; $q = 0.427$
	M8: beta & ω>1	–7401.981	16.56^{***}	$p_0 = 0.790$; $p = 0.664$; $q = 0.732$ $p_1 = 0.210$ ($\omega_s = 1.43$)
Old World *Lycium*	M1a: nearly neutral	–2425.514		$p_0 = 0.373$ ($\omega_0 = 0.103$); $p_1 = 0.627$ ($\omega_1 = 1$)
	M2a: positive selection	–2401.474	48.08^{***} *	$p_0 = 0.295$ ($\omega_0 = 0.105$); p_1=0.506 ($\omega_1 = 1$) $p_2 = 0.199$ ($\omega_2 = 4.71$)
	M7: beta	–2429.138		$p = 0.333$; $q = 0.195$
	M8: beta & ω>1	–2403.221	51.83^{***} *	$p_0 = 0.780$; $p = 0.332$; $q = 0.190$ $p_1 = 0.220$ ($\omega_s = 4.43$)

$^{****}P < 0.0001$, df=2, $^{***}P < 0.001$, df=2

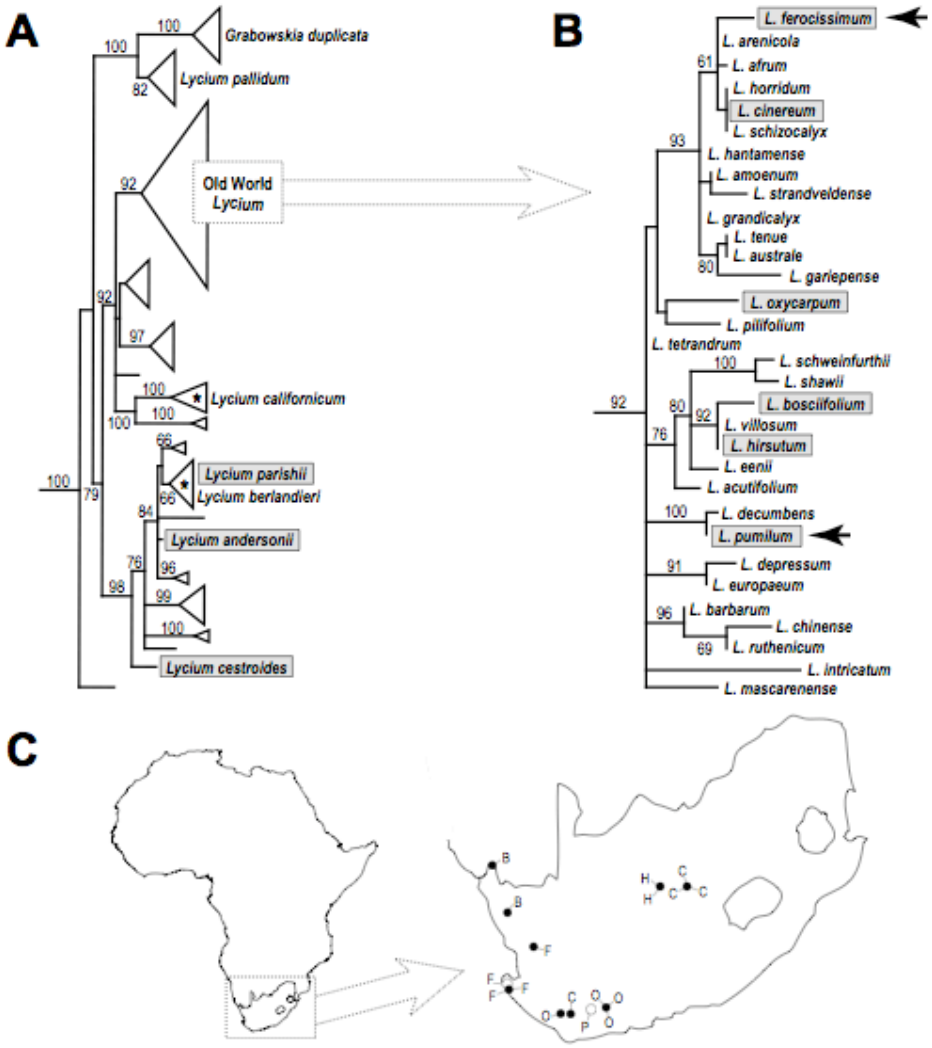

Figure 1.1. Schematic illustrating evolutionary relationships among members of tribe Lycieae (A) and Old World species of *Lycium* (B) from Levin et al. (2007); ML bootstrap values greater than 60% are indicated. Taxon names in (A) are shown only for species known to be self-incompatible (*L. andersonii*, Richman 2000; *L. cestroides*, Aguilar and Bernardello 2001; *Grabowskia duplicata*, Bianchi et al. 2000; *L. berlandieri*, *L. pallidum*, and *L. parishii*, Miller and Venable 2002; and a diploid population of *L. californicum*, Yeung et al. 2005 and JR Kohn, Univ. California, San Diego, pers. comm.). Self-compatibility is known only from two related species in one clade (*L. exsertum* and *L. fremontii*) and a polyploid population of *L. californicum* (Miller and Venable 2002), the positions of which are indicated with asterisks in (A). Taxa included in the present study are in shaded boxes, and black arrows indicate African species for which controlled pollinations were done. Maps (C) show the geographic distributions of South African species and populations in this study; letters are the first letter of the species name (see B), and locations are represented by either open circles (populations in which controlled pollinations were carried out) or closed circles (sampling localities for individuals used in *S-RNase* studies).

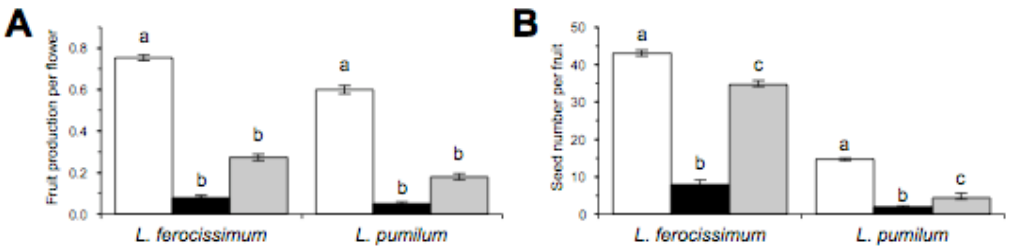

Figure 1.2. Means (±1 SE) for fruit production per flower (A) and seed number per fruit (B) in *Lycium ferocissimum* and *L. pumilum* following either outcross (open bars) or self (closed bars) pollination or in unmanipulated controls (shaded bars). For within-species comparisons of fruit or seed production, means sharing the same superscript do not differ significantly.

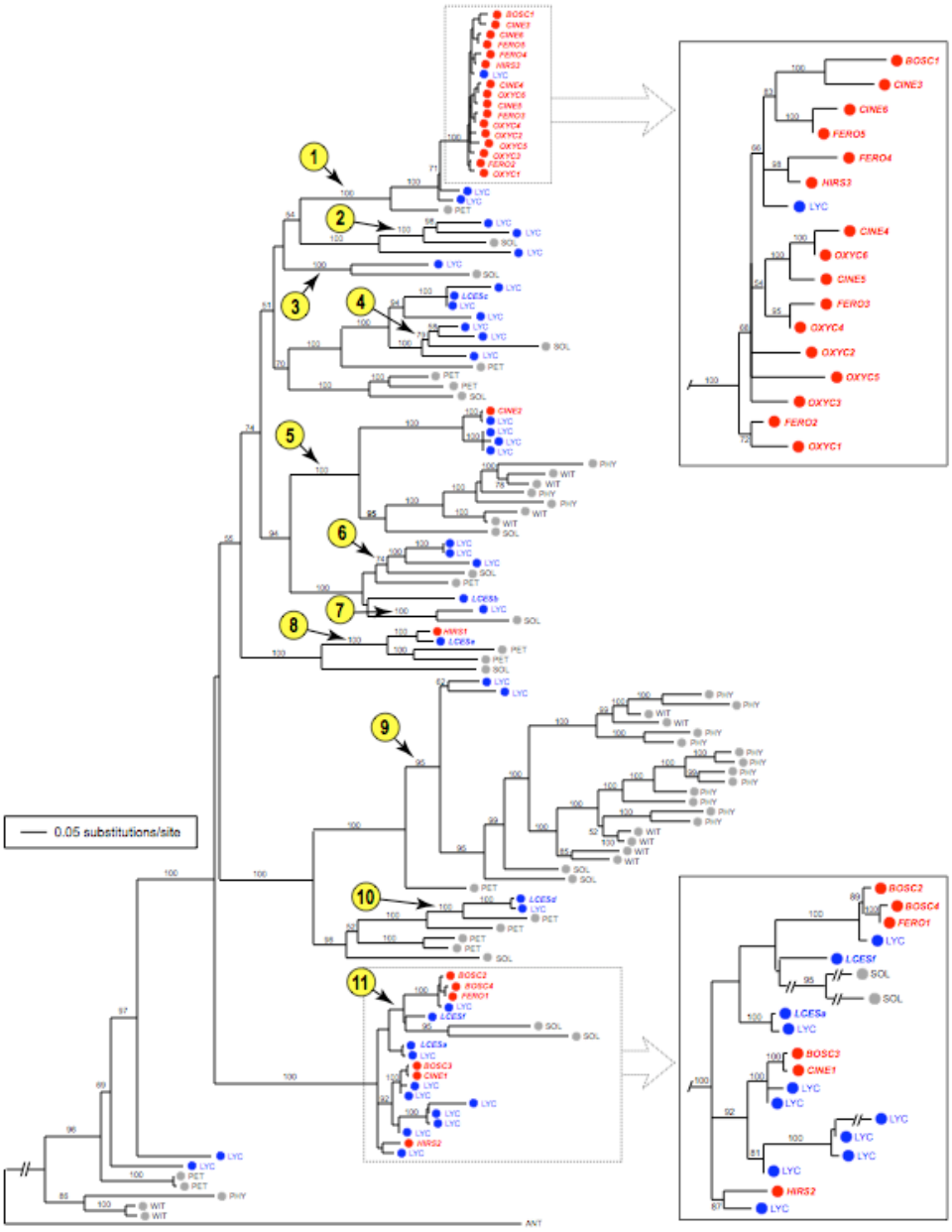

Figure 1.3. Relationships among S-RNases isolated from Old World (red circles) and New World (blue circles) *Lycium* species, and four additional genera in Solanaceae. The single tree from an ML analysis of the 120 S-RNase data set using *Antirrhinum* (ANT) as an outgroup is shown with Bayesian posterior probabilities. Trans-generic lineages for *Lycium* are indicated using circled numbers. Sequences generated in the present study are in boldface, italic type and are labeled with their species and putative allele identity following Table 1.1. Previously obtained New World *Lycium* alleles from Richman (2000) and Savage and Miller (2006) are labeled LYC. S-RNases from *Petunia* (PET), *Physalis* (PHY), *Solanum* (SOL), and *Witheringia* (WIT) are from Genbank (see Materials and Methods).

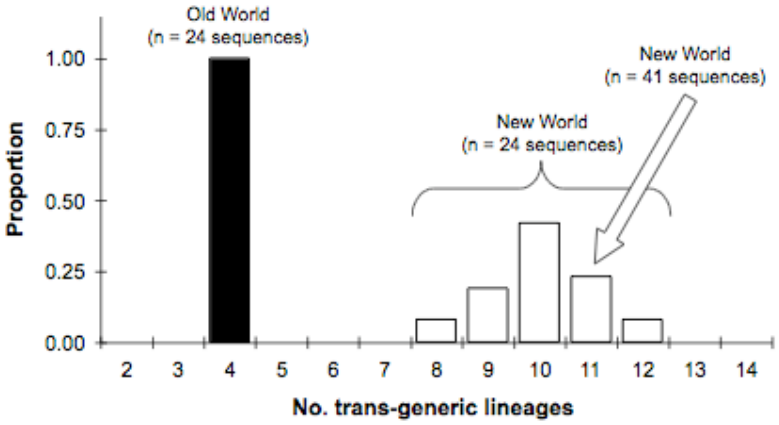

Figure 1.4. The distribution of the number of trans-generic lineages as a proportion of the topologies inferred from 100 re-sampled data sets. Each re-sampled data set included all 24 Old World sequences and a partial sample (*n*=24) of randomly chosen New World S-RNases. Open bars are the number of TGLs for New World sequences and the closed bar is the number recovered for Old World sequences in all re-sampled data sets. The open arrow indicates the number of TGLs recovered for New World S-RNases when all 41 New World S-RNases were included (see Fig. 1.3).

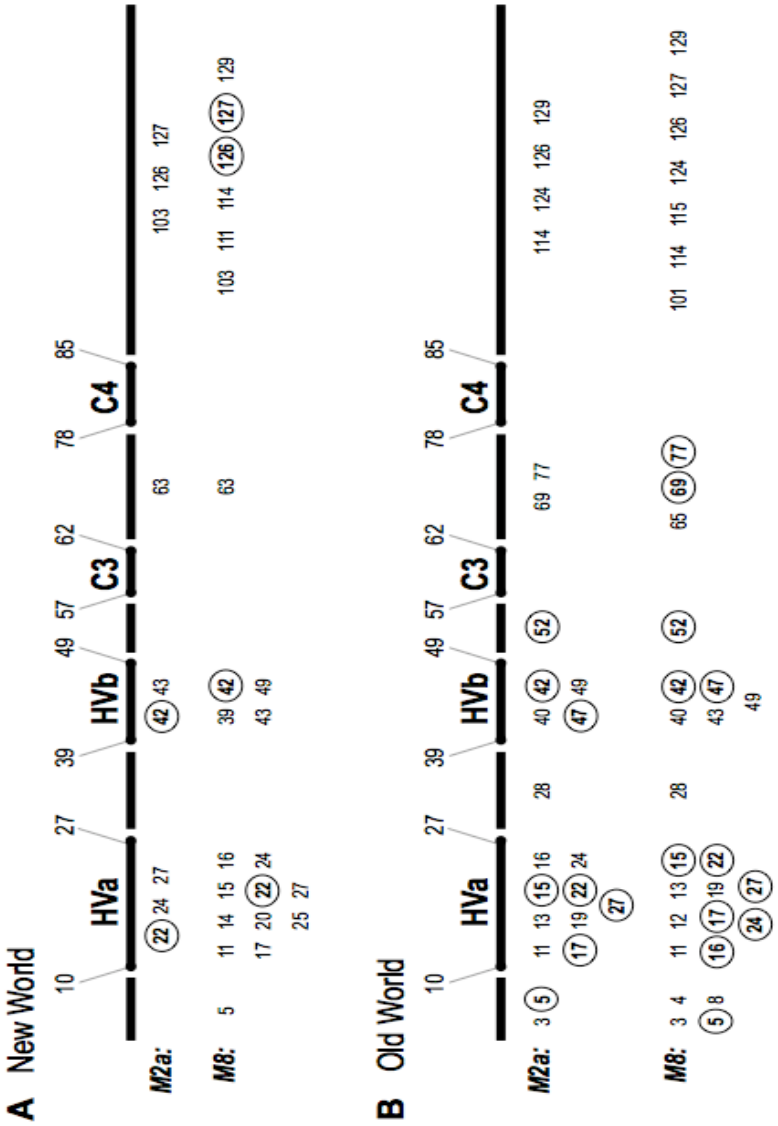

Figure 1.5. Site-specific selection at the *S-RNase* locus in New World (A) and Old World (B) *Lycium*. The sequenced region of the *S-RNase* gene is drawn to scale, and the hypervariable (HVa and HVb) and conserved regions (C3 and C4) identified previously by Ioerger et al. (1991) are indicated using numbers above the diagrams. Numbers below the diagrams are those amino acid sites inferred to be under positive selection (models M2a and M8). Circled numbers are sites with a posterior probability greater than or equal to 0.95, whereas all other positions had $0.50 \leq$ probabilities < 0.95.

CHAPTER 2

ALLELIC DIVERSITY AND SEQUENCE SIMILARITY AMONG S-RNASE ALLELES FROM NINE SPECIES OF OLD WORLD *LYCIUM*

2.1 Introduction

2.1.1 Gametophytic Self-Incompatibility

Gametophytic self-incompatibility (GSI) is a genetically controlled outcrossing mechanism that prevents self-fertilization in plants. It has evolved to inhibit inbreeding depression and promote outcrossing. In GSI, genes at a large single locus, the *S*-locus, control the self-incompatibility response. Single locus gametophytic self-incompatibility has been documented in at least 18 plant families (Mable 2004, supplementary Table 1) and is well characterized in the Solanaceae, Rosaceae, and Plantaginaceae families (Lawrence 2000; Castric and Vekemans 2004; Mable 2004).

In S-RNase based gametophytic self-incompatibility, two genes at the S-locus determine the recognition specificities. The first gene is called *S-RNase* (Fig. 2.1) and is expressed in the style of mature flowers. The second, more recently described gene is expressed in the pollen grain and has been called *SLF* (Wang et al. 2004, Sijacic et al. 2004) in *Petunia* (Solanaceae) or *SFB* (Ushijima et al. 2001; Entani et al. 2003) in *Prunus* (Rosaceae). Pollen tube growth is aborted when the allele specificities of these two genes match; pollen tube growth continues when the allele specificities are different.

The *S-RNase* gene is well characterized and encodes a functional style-specific RNase (McClure et al. 1989) that is necessary for the rejection of self-pollen (Murfett et al. 1994), though successful rejection of self pollen also relies on non-S specific factors (Goldraij et al. 2006; McClure 2006). The *S-RNase* gene is comprised of two hypervariable regions (HVa and HVb) and five conserved regions (C1-C5), with an intron nested in hypervariable region HVa (Fig 2.1; Ioerger et al. 1991). The hypervariable regions are hydrophilic portions of the *S-RNase* gene (Ioerger et al. 1991) that play a role in allele recognition (Matton et al. 1997). They lie on the outer surface of the S-RNase protein where they can interact with pollen *S*-specificities (Ida et al. 2001). Less is known about the *SLF* (*S*-locus F-box) gene. It contains an F-box domain (a site of protein-protein interaction) that plays a role in the recognition of non-self S-RNases (Qiao et al. 2004; Lai et al. 2002). How the *S-RNase* and *SLF* genes interact with each other on a molecular level is an area of research being actively pursued (McClure and Tong 2006; Goldraij et al. 2006; Wang et al. 2003; Kao and Tsukamoto 2004).

A central feature of gametophytic self-incompatibility systems is the maintenance of high numbers of alleles within populations (Wright 1939). For example, Lawrence (2000) and Castric and Vekemans (2004) have reviewed allele number in plants with GSI and found high numbers of alleles with a large amount of sequence divergence in population surveys. Such polymorphism is likely maintained by strong negative frequency dependent selection, which causes the retention of alleles in GSI populations. Rare alleles have a mating advantage because they are compatible with any S-genotype other than themselves in a population. This rare allele advantage makes them less likely to be lost from a population. Likewise, new alleles that arise through mutation or that

enter a population by gene flow also have a mating advantage since they are by definition rare. Through time, equal numbers of alleles will be maintained as rare allele specificities become more common. In populations experiencing negative frequency dependent selection, large numbers of alleles are maintained because they are rarely lost and selection is strong for novel specificities. Large numbers of alleles are selected for to ensure that individuals will become pollinated by an allele that is not their own (Wright 1939).

A second and much longer-term consequence of negative frequency dependence is the maintenance of alleles in descendent taxa through speciation events. This shared S-RNase polymorphism among species and genera is termed trans-generic polymorphism (Ioerger et al. 1990). This implies that *S-RNase* alleles predate the diversification of the taxa they reside in (Ioerger et al. 1990, Richman and Kohn 1999; Igic 2006; Savage and Miller 2006). *S-RNase* alleles that are from two different genera and that also share recent ancestry are said to be in the same trans-generic lineage.

Extensive numbers of trans-generic lineages have been documented at the *S-RNase* locus in the Solanaceae. Population surveys of *S-RNase* alleles in the Solanaceae have found varying numbers of trans-generic lineages (TGLs) for alleles in different species [17 TGLS in *Solanum chilense* (Igic et al. 2007), eight TGLs in *S. carolinense* (Richman and Kohn 2000), 12 TGLs in *Lycium andersonii* (Richman and Kohn 2000) and 10 TGLs in *L. parishii* (Savage and Miller 2006)]. The presence of closely related *S-RNase* alleles across six genera in Solanaceae (Richman and Kohn 2000; Igic et al. 2004) has been interpreted as evidence of the uninterrupted maintenance of GSI over the long term (Igic et al. 2006).

The pattern and number of trans-generic lineages has been useful for reconstruction of demographic events in the evolutionary history of Solanaceae. In contrast to the large numbers of trans-generic lineages found in most Solanaceae, *S-RNase* alleles isolated from *Physalis* and *Witheringia* are found in a reduced number of TGLs. Evidence for a historical *S-RNase* allele bottleneck in the ancestor of *Physalis crassifolia*, *P. cinerascens*, *P. longifolia*, *Witheringia maculata* and *W. solanacea* has been found through the analysis of trans-generic lineages (Richman et al. 1996; Richman and Kohn 1999, 2000; Lu 2001; Igic et al. 2004; Stone and Pierce 2005). *Physalis* and *Witheringia*, which are sister species, form a monophyletic clade that is about 14 million years old (Paape et al. in press). Alleles in *P. crassifolia* are numerous and closely related, whereas *P. cinerascens* alleles are less numerous and more highly diverse than *P. crassifolia* alleles. Although there is a discrepancy between the allele number and diversity in these two species, their alleles fall into the same three TGLs (Richman et al. 1996; Richman and Kohn 1999, 2000; Richman 2000; Igic et al. 2004). Expanded analyses of *S-RNase* diversity including additional species (*P. longifolia, W. maculata* and *W. solanacea*) also support the bottleneck hypothesis as all recovered alleles are found in the same three TGLs (Lu 2001; Stone and Pierce 2005). The restricted number of TGLs found in *Physalis* and *Witheringia* suggest that these genera have lost trans-generic lineages in a bottleneck event (Richman et al. 1996; Richman and Kohn 1999, 2000; Richman 2000; Lu 2001; Igic et al. 2004; Stone and Pierce 2005). It also suggests that the bottleneck that led to the same limited number of TGLs in *Physalis* and *Witheringia* occurred in a common ancestor of these two genera (Stone and Pierce 2005).

Patterns of *S-RNase* allele diversification and molecular evolution (such as synonymous and non-synonymous substitution rates) give insight into the after affects of allele bottlenecks. An excess of non-synonymous mutations in the hypervariable regions of the *S-RNase* gene (compared to other regions of the gene) in *Witheringia* suggest that strong positive selection is driving the re-diversification of these *S-RNase* alleles into new specificities (Richman and Kohn 2000; Stone and Pierce 2005). Intraspecific comparisons of *Witheringia S-RNase* alleles show higher nonsynonymous substitution rates than interspecific comparisons of alleles, suggesting the re-diversification of *S-RNase* alleles post-speciation (Strong and Pierce 2005). In contrast, inter and intraspecific non-synonymous/synonymous ratios between *Physalis crassifolia* and *P. cinerascens* are similar, suggesting that alleles in these genera diversified in a common ancestor following the initial bottleneck, but prior to the divergence of these taxa (Richman 2000; Richman and Kohn 2000; Stone and Pierce 2005).

2.1.2 Study System

The genus *Lycium* (Solanaceae) contains approximately 80 species found throughout the world, with centers of diversity in South America (ca. 30 species; Bernardello 1986; Hitchcock 1932; Levin and Miller 2005), North America (21 species; Hitchcock 1932; Chiang-Cabrera 1981; Miller 2002) and southern Africa (26 species; Venter 2000, 2007). Species in this genus are thorny shrubs that are found in dry and often saline, coastal environments. *Lycium* typically has multi-seeded red, and occasionally yellow or black, fleshy berries (Levin and Miller 2005; Levin et al. 2007) that are attractive to birds and may help explain its leading to its cosmopolitan distribution. A number of species in this genus are self-incompatible (Miller and Venable 2000, 2002; Richman and Kohn 2000;

Aguilar and Bernadello 2001; Savage and Miller 2006; Miller et al. in press). Ten (three in North America and seven in southern Africa) *Lycium* species are gender dimorphic, while the remaining species are hermaphroditic. Populations of the North American dimorphic species are morphologically gynodioecious (containing separate female and hermaphroditic plants), but hermaphrodites are functionally male (Miller and Venable 2002) whereas African dimorphic taxa have separate male and female plants (Venter 2000, 2007).

African *Lycium* are found in both the northern and southern regions of the continent, but species diversity is considerably higher in the south. Three species, *Lycium schweinfurthii*, *L. europaeum* and *L. intricatum*, are found along the northwest and Mediterranean coasts of northern Africa. In contrast, twenty-three species of *Lycium* are present in southern Africa, occupying the countries of Namibia, Botswana, Zimbabwe, Mozambique, Angola and extensively in South Africa (Venter 2000). One species, *L. shawii*, has a distribution along the eastern coast of Africa from South Africa to Egypt (Venter 2000). This is the only species that spans the length of the continent and unites the southern *Lycium* species with those in the north. Old World *Lycium* species (including three East Asian *Lycium* sampled to date) as a whole form a strongly supported monophyletic group (Levin and Miller 2005; Levin et al. 2007); however, it remains unclear whether northern African *Lycium* are more closely related to southern African or to East Asian species.

2.1.3 Self-incompatibility in Old World *Lycium*

Baker (1955, 1967) was the first to assert that it is more likely for a single propagule of a self-compatible individual to start a sexually reproducing colony than it is for two self-

incompatible propagules by chance to be in the same vicinity at the right time to start a sexually reproducing colony. He supported this assertion with results from studies chronicling the existence of a disproportionate number of self-compatible species on oceanic islands compared to self-incompatible individuals.

Phylogenetic evidence strongly suggests *Lycium* dispersed from the New World to the Old World (Africa and Asia) a single time approximately 2.13-6.79 million years ago (Levin and Miller 2005; Shak 2006). In chapter one, controlled pollinations of two species of Old World *Lycium* were performed to test for the presence of self-compatibility in the Old World. These pollinations demonstrate that the two studied species, *Lycium ferocissimum* and *L. pumilum*, were strongly self-incompatible.

Species that have maintained self-incompatibility and that have undergone a recent S-RNase bottleneck contain *S-RNase* sequences that are under strong positive selection to diversify, thereby increasing the number of their alleles (Richman and Kohn 1996, Miller et al. in press). This has been found in post-bottleneck *Witheringia* alleles (Stone and Pierce 2005, Richman and Kohn 2000). Previous selection analyses on amino acid sequences from Old World and New World *Lycium* S-RNases suggest that positive selection is stronger and acts on a larger number of amino acid sites in Old World sequences compared to New World sequences, suggesting Old World sequences are re-diversifying (see chapter one). These data are consistent with the documentation of self-incompatibility in *Lycium pumilum* and *L. ferocissimum* (see chapter one).

When self-incompatibility breaks down, *S-RNase* polymorphism is lost because the absence of negative frequency dependence no longer causes selection for new allele specificities or retains alleles in a population. Because of this, neutral mutations

accumulate and alleles are subject to genetic drift. Igic et al. (2006) have argued that following the breakdown of self-incompatibility, the loss of *S-RNase* polymorphism leads to the monophyly of S-RNases (Kondo et al. 2002). Conversely, the presence of shared ancestral polymorphism among species can be used as an indication that self-incompatibility is the ancestral state of taxa. Shared allele polymorphism of Old World *Lycium S-RNase* alleles would be an indication that self-incompatibility was retained in Old World *Lycium* despite long-distance dispersal.

2.1.4 Current Study

I previously isolated 24 *S-RNase* alleles from 15 Old World *Lycium* individuals from five different species (*Lycium bosciifolium, L. cinereum, L. ferocissimum, L. hirsutum* and *L. oxycarpum*) (chapter one). These alleles, along with 44 alleles from previously published studies of New World *Lycium* species were used to construct a maximum likelihood S-RNase genealogy. Four Old World and 11 New World *Lycium* trans-generic lineages were identified, indicating that Old World *Lycium* have a restricted number of trans-generic lineages. I interpret these data as evidence of a bottleneck at the *S-RNase* locus, an interpretation consistent with phylogenetic studies indicating a single dispersal event of the genus to the Old World (Levin and Miller 2005; Levin et al. 2007).

In chapter one, the number of trans-generic lineages present in the Old World was estimated by the construction of an *S-RNase* allele genealogy that included *S-RNase* alleles from New and Old World *Lycium* as well as alleles from other species in the Solanaceae. Twenty-four Old World and 44 New World *Lycium* alleles were used to construct the genealogy. Because there were a greater number of New World alleles in the genealogy, it could potentially cause a bias in the number of New and Old World

lineages recovered. An attempt was made to remedy this imbalance by randomly re-sampling 24 New World alleles (the same number of alleles found in the Old World species) and using them to construct genealogies and count TGLs with the Old World *Lycium* and larger Solanaceae data set. Following 100 re-sampled datasets, eight-twelve New World and four Old World *Lycium* TGLs were recovered.

A second approach to equalizing New and Old World *Lycium* allele sampling is to increase sampling from among the Old World species. The original dataset (chapter one) of Old World *S-RNase* alleles was isolated from 15 individuals from five species. In this study, I expanded my sampling to include an additional 12 individuals from 7 species (*Lycium afrum*, *L. ferocissimum*, *L. gariepense*, *L. hantamense*, *L. hirsutum*, *L. oxycarpum* and *L. tenue*). Overall, the Old World S-RNase dataset was expanded by ca. 45% (from 15 to 27 individuals from 5 to 9 species). Thus, I isolated an additional 12 alleles from seven species, four of which (*Lycium afrum*, *L. gariepense*, *L. hantamense* and *L. tenue*) were not represented in the previously published data set (Miller et al. in press), bringing the total number of Old World alleles to 36. With an increase in the number of Old World *S-RNase* alleles, I am testing whether additional TGLs will be recovered and if the diversity will be increased within existing ones. I used PAML (Phylogenetic Analysis using Maximum Likelihood; Yang 1997) to estimate site-specific d_N/d_S ratios for the *S-RNase* gene for this larger data set. This was done to reexamine selection in both New and Old World *Lycium*. Because there is greater selection for the re-diversification of alleles post bottleneck (Richman et al. 1996; Richman and Kohn 1999, 2000; Stone and Pierce 2005), I expect to find a larger number of positively selected sites in Old World *Lycium* than in the New World.

The dispersal of *Lycium* to the Old World is likely quite recent (ca. 2.13-6.79 mya; Shak 2006), which means that Old World *Lycium* are very closely related to one another. Few studies have examined *S-RNase* alleles in closely related species, except for the re-diversification of *S-RNase* alleles post-bottleneck in *Physalis* and *Witheringia* (Richman et al. 1996). The bottleneck in *S-RNase* alleles following the recent dispersal of *Lycium* to the Old World provides an excellent opportunity to study the *S-RNase* locus in species that have diverged recently. Given that the radiation of Old World *S-RNase* alleles occurred after the dispersal of *Lycium* to the Old World, then I expect to recover Old World *S-RNase* sequences that have diversified from the same limited number of ancestral alleles. Selection for the diversification of alleles should lead to differences in these *S-RNase* alleles, albeit minimal differences since Old World species are young. I also expect to see higher pair-wise amino acid similarity among Old World *S-RNase* alleles than among New World alleles. Studying the re-diversification of Old World *Lycium S-RNase* alleles will provide insight into patterns of within lineage diversification, such as the number of new alleles that can be derived from a single trans-generic lineage and still persist within a population.

Interestingly, seven partial *S-RNase* sequences (ca. 380 bp, 61% of the *S-RNase* gene) in this study had identical coding sequences that were shared across different Old World *Lycium* species. To investigate these sequences further, I designed allele specific primers to isolate the coding region and intron from genomic DNA of individuals that shared these identical partial *S-RNase* alleles. Introns accumulate mutations neutrally through time. Identical introns in different species could indicate that not enough time has passed for *S-RNase* alleles and introns to diversify post-speciation or that *S-RNase* alleles were

inherited from a process such as gene-flow. A phylogeny of *Lycium* species that have identical partial alleles was also created to determine if *S-RNase* alleles are more likely to be identical among closely related species than among distantly related ones.

2.2 Methods

2.2.1 Isolation of *S-RNase* Alleles from Style Tissue

Styles for this study were collected from 12 individual *Lycium* plants in southern Africa (Fig. 2.2) from July to August 2005 and from August to September 2006 and preserved immediately in RNAlater® (Ambion, Inc., Austin, TX). Styles were collected from *L. afrum, L. gariepense, L. hantamense, L. hirsutum, L. oxycarpum* and *L. tenue*. It is not possible to use universal primers (i.e. primers that amplify all alleles) to extract *S-RNase* sequences from genomic DNA because *S-RNase* alleles are so highly divergent from one another. Instead, the *S-RNase* gene is isolated from RNA style tissue using degenerate primers.

Total RNA was isolated from 7-10 styles of an individual plant using a Qiagen RNeasy Plant Mini Kit (Qiagen, Inc., Valencia, CA) and cDNA was synthesized from this RNA using a Novagen First Strand cDNA Synthesis Kit (EMD Biosciences, Inc., Madison, WI). Degenerate primers PR1 and PR3 (Richman et al. 1995) were used during PCR to amplify the *S-RNase* gene from conserved regions C2 to C5 (Fig. 2.1). For a 50 µl PCR reaction, 5 µl of cDNA product, 1 µl of 10mM dNTPs, 5 µl 10X Taq buffer2 µl of 50ng/µl primers PR1 and PR3 and 0.4 µl of Taq DNA polymerase were used. The reaction was amplified with the following PCR conditions: 94°C (5 min.); 40 cycles at 94°C (15 sec.), 45°C (1 min.) and 72°C (1 min.); 72°C (7 min.). Amplification products

were cleaned using the Qiagen Qiaquick PCR Purification Kit (Qiagen, Inc., Valencia, CA) and cloned using the Novagen pT7Blue Perfectly Blunt Cloning Kit (EMD Biosciences, Inc., Madison, WI) to separate different alleles within individuals. To isolate DNA from bacterial colonies, individual white colonies and colonies with a faint blue dot were lifted with a sterile pipette tip and suspended in 50 µl of water. Bacterial suspensions were placed in a 99°C water bath for 5 minutes and then centrifuged for 1 minute at 13,000 RPM. The supernatant, containing bacterial plasmid DNA, was transferred to a new microcentrifuge tube and stored at -20°C.

To screen colonies for the presence of *S-RNase* sequences, colony DNA was amplified in a 25 µl PCR reaction with primers PR1 and PR3, 5 µl of colony DNA and the PCR conditions and reagent concentrations used to amplify cDNA, but scaled to a 25 µl PCR reaction. Amplification products from this PCR were cut with various restriction enzymes and restriction fragment length polymorphisms (RFLPs) were used to identify different alleles. DNA with different RFLP banding patterns were interpreted as having different alleles. Colonies were selected for sequencing based on differences in RFLP patterns and cleaned using the Qiagen PCR Purification Kit (Qiagen, Inc., Valencia, CA). At times, colonies did not show different banding patterns because their sequences were so similar. In this case, additional colonies were chosen at random for sequencing. A final PCR was performed on colony DNA using vector specific primers U-19 (5'-GTT TTC CCA GTC ACG ACG T-3') and R-20 (5'-CAG CTA TGA CCA TGA TTA CG-3') with the following PCR conditions: 95°C (5 min.); 2 cycles at 94°C (30 sec.), 55°C (1 min.), 72°C (2 min.); 2 cycles at 94°C (30 sec.), 54°C (1 min.), 72°C (2 min.); 2 cycles at 94°C (30 sec.), 53°C (1 min.), 72°C (2 min.); 30 cycles at 94°C (30 sec.), 52°C (1 min.),

72°C (2 min.); 72°C (5 min.). For a 50 µl PCR reaction, 10 µl of cloned genomic DNA, 1 µl of 10mM dNTPs, 5 µl 10X Taq buffer, 1 µl each of 5 pmol/µl primers R-20 and U-19 and 0.25 µl Taq DNA polymerase were used. Sequencing was carried out on an Applied Biosystems Automated 3730 DNA Analyzer by the Biotechnology Resource Center at Cornell University (Ithaca, NY).

2.2.2 Genomic PCR

Of the 12 *Lycium* individuals genotyped in this study and the 15 individuals genotyped in Miller et al. (in press), six pairs and one triad of partial alleles from different species were identical at the nucleotide level. Thus, to further explore these similarities, allele-specific primers were designed and the *S-RNase* gene was amplified using genomic DNA as a template. In particular, I was able to amplify the single intron (Ioerger et al. 1991) present in the *S-RNase* gene to determine if the alleles were identical in both the coding and non-coding regions.

Leaf tissue was collected in silica desiccant for seven (six species) of the 15 (seven species) southern African *Lycium* individuals that shared identical partial sequences. DNA was extracted from leaf material using a Qiagen DNeasy Plant Mini kit (Qiagen, Inc., Valencia, CA). Because two alleles are found in diploid individuals, I designed primers for unique regions of each allele. One degenerate primer and eight allele specific primers (Table 2.1) were used to isolate *S-RNase* sequences from genomic DNA of individuals with shared alleles. This was done to confirm that the same partial allele sequence, including the intron found in hypervariable region HVa (Ioerger et al. 1991), was shared among different species. For a 50 µl PCR reaction, 2 µl of DNA, 5 µl of 10X Taq buffer, 1.25 µl 10µM dNTPs, 2 µl 25mM MgCl2, 1 µl 10uM Primers and 0.25 µl

Taq DNA polymerase were used. PCR conditions were the following: 95°C (5 min.); 2 cycles at 94°C (30 sec.), 55°C (1 min.), 72°C (2 min.); 2 cycles at 94°C (30 sec.), 54°C (1 min.), 72°C (2 min.); 2 cycles at 94°C (30 sec.), 53°C (1 min.), 72°C (2 min.); 2 cycles at 94°C (30 sec.), 52°C (1 min.), 72°C (2 min.); 2 cycles at 94°C (30 sec.), 51°C (1 min.), 72°C (2 min.); 2 cycles at 94°C (30 sec.), 50°C (1 min.), 72°C (2 min.); 28 cycles at 94°C (30 sec.), 49°C (1 min.), 72°C (2 min.); 72°C (10 min.). Primer pairs and the alleles they amplified are listed in Table 2.2.

Styles were not collected for three *Lycium hirsutum* individuals; instead, only genomic material was collected. For these individuals, DNA was screened for the presence of the HIRS4 and HIRS6 alleles using primers ZA-SA-f and ZA-SA-r. Genomic DNA was cloned following the above protocol, and bacteria colonies were screened for alleles using primers ZA-SA-f and ZA-SA-r instead of primers PR1 and PR3. Cloned DNA containing putative *S-RNase* alleles were amplified with primers R20 and U-19 before being sent to sequencing. Amplification products were sequenced at the Biotechnology Resource Center at Cornell University (Ithaca, NY). Individuals that differed by one base pair in their intron were amplified an additional time and sequenced to confirm the one bp difference.

2.2.3 Genealogy and Phylogeny Construction Using Maximum Likelihood

2.2.3.1 *S-RNase* Allele Genealogy

One hundred and thirty three *S-RNase* sequences were aligned by eye in Sequence Alignment Editor (Se-Al) version 2.0a11 (Rambaut 1996). Amino acid sequences were aligned by eye using Ioerger et al. (1991) as a reference. *S-RNase* sequences from both

Old World (36 sequences) and New World *Lycium* (44 sequences), as well as from other members of the Solanaceae (50 sequences) and from *Antirrhinum* (outgroup; 3 sequences) were used in the alignment (Table 2.3).

An *S-RNase* genealogy using a maximum likelihood (ML) algorithm was constructed to infer the evolutionary histories of *S-RNase* lineages. The Akaike information criterion (AIC) as implemented by Modeltest version 3.7 (Posada and Crandall 1998) was used to determine the most appropriate model of sequence evolution for the ML analysis. The best-fit model selected by AIC was TVM+I+G (nucleotide frequencies: A=0.37750, C=0.19370, G=0.2240, T=0.22640; substitution rate matrix: A-C=1.636, A-G=2.7756, A-T=1.2826, C-G=1.9095, C-T=2.7756, G-T=1.0000; assumed proportion of invariable sites: 0.0482; gamma distribution shape parameter: 1.6787). This model was used in a PAUP* (Phylogenetic Analysis Using Parsimony) version 4.0b10 (Swofford 2002) ML analysis with a single neighbor-joining tree as a starting topology. PAUP* ML settings included the heuristic search option, tree-bisection reconnection (TBR) branch-swapping and the MulTrees options in effect. Three *Antirrhinum hispanicum S-RNase* sequences were defined as the outgroups. A PAUP* ML nonparametric bootstrap analysis was conducted on the Condor (Condor Project, 2005) computer cluster at Amherst College. For this analysis 100 heuristic bootstrap replicates were run, each using three random addition sequence replicates and NNI (nearest neighbor interchange) branch-swapping; the MulTrees option was not in effect. The analysis was divided into 100 pieces and RepMaker (Wiglenbusch 2003) was used to partition these jobs on the cluster. Old and New World *Lycium* trans-generic lineages (TGLs) identified from the *S-RNase* allele ML

genealogy. TGLs were defined as the youngest node on a genealogy that includes a *Lycium* sequence and a sequence from another genus.

2.2.3.2 *Lycium* Phylogeny

Using data from the nuclear granule-bound starch synthase GBSSI (*waxy*) gene and chloroplast regions *trnT-trnL* and *trnD-trnT* (Shak 2006), I reconstructed phylogenetic relationships among the Old World *Lycium* sampled in the present study. Sequences from 14 Old World *Lycium*, one New World *Lycium* and two New World *Nolana* (outgroup) individuals were aligned by eye using Se-Al version 2.0a11 (Rambaut 2002). The AIC model selected by Modeltest version 3.7 (Posada and Crandall 1998) was TVM+G (nucleotide frequencies: A=0.3089, C=0.1703, T=0.1868; substitution rate matrix: A-C=0.5987, A-G=2.0764, A-T=0.5208, C-G=1.5742, C-T=2.0764, G-T=1.0000; assumed proportion of invariable sites: 0; gamma distribution shape parameter: 0.3668). This model was used to create a ML Phylogeny in PAUP* version 4.0b10 (Swofford 2002) using a single neighbor-joining tree was used as a starting topology. PAUP* ML settings included the heuristic search option and tree-bisection reconnection (TBR) branch-swapping; the MulTrees option was in effect. ML bootstrap values were calculated in PAUP* by running two hundred bootstrap replicates with the PAUP* default options.

2.2.4 Selection and Diversity at the *S-RNase* Gene

Selection on the *S-RNase* gene was estimated using the *codeml* package in Phylogenetic Analysis by maximum likelihood (PAML; Yang 1997). *Codeml* was used to find site specific d_N/d_S (ratio of non-synonymous substitutions per non-synonymous

site to the ratio of synonymous substitutions per synonymous site) rate ratios (ω). Analyses for New and Old World *Lycium* were performed separately. Additionally, the average pairwise amino acid distances at the *S-RNase* gene was calculated using PAUP* (Swofford 2002) for both New and Old World *Lycium*.

The recommended procedure in PAML to test for positive selection is to compare nested sets of different models of selection. Four models of selection were tested in PAML using likelihood ratio tests (LRTs). Nearly neutral models of evolution [models M1a and M7 (beta)] were compared to models that incorporated positive selection [models M2a and M8 (beta&ω)] to determine how sequence sites were evolving. In the nearly neutral M1a model, amino acid sites were tested for purifying selection ($0 < \omega_0 < 1$) and neutral or nearly neutral evolution ($\omega_1 = 1$). This model was compared to model M2a, which tested for neutral or nearly neutral evolution, purifying selection and positive ($\omega_2 > 1$) selection. Nearly neutral model M7 (beta) was compared to model M8 (beta&ω), which tests for positive selection ($\omega > 1$). Models M7 and M8 assume a β-distribution to account for purifying selection ($0 < \omega < 1$). LRTs were used to compare null model M1a (neutral and purifying selection) against model M2a (positive selection) and model M7 (neutral evolution and purifying selection) against model M8 (positive selection). A chi square test with two degrees of freedom was used to determine whether a model incorporating positive selection (models M2a and M8) fit the data significantly better than a simpler model (models M1a and M7). Posterior probabilities for each amino acid site were calculated using an empirical Bayes method to determine which sites are under neutral evolution, purifying selection or positive selection.

2.3 Results

2.3.1 Isolation of *S-RNase* Alleles

In this study, a total of 12 partial *S-RNase* allele sequences were isolated from stylar RNA from 12 individuals and seven species of Old World *Lycium* (Table 2.4). These partial allele sequences were from seven different species (*L. afrum, L. ferocissimum, L. gariepense, L. hantamense, L. hirsutum, L. oxycarpum* and *L. tenue*) and ranged from 376-384 bp long. Identical allele nucleotide coding regions were found in *Lycium hirsutum* and *L. ferocissimum* (HIRS6/ FERO3), *L. ferocissimum* and *L. tenue* (FERO1/TENU2), *L. oxycarpum* and *L. cinereum* (OXYC7/CINE1), *L. oxycarpum* and *L. hirsutum* (OXYC1/HIRS4), *L. oxycarpum* and *L. ferocissimum* (OXYC5/FERO6), *L. cinereum* and *L. hirsutum* (CINE5/HIRS5) and *L. hantamense*, *L. bosciifolium* and *L. oxycarpum* (HANT2/BOSC2/OXYC8) (Table 2.4).

I used genomic DNA to isolate allele sequences from individuals carrying putatively identical *S-RNase* alleles BOSC2/HANT2/OXYC8, FERO1/TENU2 and HIRS4/OXYC1 (Fig. 2.3). These allele sequences contained coding and non-coding regions of the *S-RNase* gene and were ca. 433-471 bp long. The partial coding region and intron of allele BOSC2/HANT2/OXYC8 was identical in *Lycium bosciifolium*, *L. hantamense* and *L. oxycarpum* individuals that carried this allele. The partial coding region and intron of the putatively shared allele HIRS4/OXYC1 in *L. oxycarpum* and *L. hirsutum* were also identical. The partial coding region of *S-RNase* allele FERO1/TENU2 in *L. ferocissimum* and *L. tenue* was the same in each individual that carried the allele, but the introns had a single bp difference from one another. The second allele in each individual was also sequenced from genomic DNA, except for *L. hantamense* and *L. hirsutum*. Genomic

material was not available for the isolation of non-coding regions in individuals carrying shared *S-RNase* sequences FERO3/HIRS6, OXYC7/CINE1, CINE5/HIRS5 and OXYC5/FERO6.

2.3.2 *S-RNase* Genealogy and *Lycium* Phylogeny

Genealogical analysis indicates that *S-RNases* from *Lycium* fall into 12 different trans-generic lineages. New World *Lycium* S-RNases were present in all 12 of the TGLs whereas alleles from Old World *Lycium* were present in only four of them (Fig. 2.4). Old World *Lycium* alleles were sister to six different New World *Lycium* alleles, indicating that six different New World alleles may have been dispersed to the Old World. Twenty-four Old World *S-RNase* alleles were found in TGL one, ten in TGL four, and one each in TGLs two and three.

The Old World *Lycium* phylogeny constructed using the nuclear GBSSI gene and the noncoding chloroplast *trnD-trnT* and *trnT-trnL* regions contains a monophyletic group of Old World *Lycium* (Fig. 2.5). This monophyletic group has two sister clades, clade A and B. Clade A consists of one highly supported group (bootstrap value 96 in Fig. 2.5) that is sister to *L. oxycarpum*. Clade B has moderate support (bootstrap value 74 in Fig. 2.5) and consists of a polytomy of three individuals and a highly supported group of three individuals.

2.3.3 Selection and Diversity at the *S-RNase* Gene

The average pairwise amino acid distance for Old World *Lycium* alleles was 0.31 and for New World alleles was 0.53. Old World *Lycium bosciifolium, L. hantamense* and *L. oxycarpum* allele BOSC2/HANT2/OXYC8 had only six amino acid differences from

New World *L.* p*arishii* allele LPAR19. New World *L. parishii* allele LPAR6 also had six amino acid differences from Old World *L. bosciifolium* allele BOSC3 and the *L. cinereum* and *L. oxycarpum* allele CINE1/OXYC7. New World *L. parishii* allele LPAR11 had only three amino acid differences from Old World *L. cinereum* allele CINE2.

The *codeml* analysis executed in PAML suggests that both models of positive selection, M2a and M8, fit the sequence data better than the models of nearly neutral selection, M1a and M7, for both Old World and New World *Lycium* (Table 2.5). In a comparison of model M1a (nearly neutral selection) and model M2a (positive selection), model M2a fits the data significantly better than model M1a for Old World (LRT = 46.412, P = $8.35e^{-11}$) and New World (LRT = 14.812, P = 0.0008) *Lycium*. In a comparison of model M7 (nearly neutral selection) and model M8 (positive selection), model M8 fits the data significantly better than model M7 for Old World (LRT = 50.154, P = $1.286e^{-11}$) and New World (LRT = 16.537, P = 0.0003) *Lycium*. There were twice as many sites under positive selection for Old World *Lycium* model M2a than for New World model M2a and about one and a half times as many sites under positive selection in Old World *Lycium* model M8 than for New World *Lycium* model M8.

PAML analyses on the *S-RNase* gene suggest that both Old World and New World allele sequences are undergoing positive selection, although Old World alleles have eight times more sites under strong positive selection than New World alleles under model M2a and about four times more sites under strong positive selection under model M8. Nine sites under the M2a model and 19 sites under the M8 model were under selection for New World *Lycium*. In New World *Lycium*, one site in the M2a and four sites in the

M8 model had a greater than 95% probability of being under positive selection. Twenty-one sites for the M2a model and 30 sites for the M8 model were under positive selection for Old World *Lycium*. In Old World *Lycium* eight sites for the M2a model and eleven sites for the M8 model had a greater than 95% probability of being under positive selection. Seventeen out of the 30 sites under positive selection in Old World *Lycium* and 12 out of the19 sites under positive selection in New World *Lycium* were found in hypervariable regions HVa and HVb (Fig. 2.6). No sites under positive selection were found in conserved regions C3 or C4 for New and Old World *Lycium*. Fourteen of the same sites were undergoing positive selection in both New and Old World *Lycium*.

2.4 Discussion

2.4.1 S-RNase Genealogy

To address the issue of unequal New and Old World *Lycium* allelic sampling present in chapter one, 12 additional Old World *Lycium* partial *S-RNase* alleles were isolated and included with the 24 Old World alleles isolated in chapter one. Thirty-six New World and 44 Old World alleles, along with *S-RNase* alleles from species in the Solanaceae family, were used to create a S-RNase genealogy. The addition of 12 Old World *S-RNase* alleles extended the diversity of two out of the four TGLS (Fig. 2.4; TGLs one and four) identified previously, but did not increase the number of TGLs identified. Similarly, in other studies the inclusion of additional *Physalis* and *Witheringia S-RNase* alleles (Stone and Pierce 2005) to existing genealogies (Richman et al. 1996; Richman 2000; Richman and Kohn 1999, 2000) only increases the diversity of alleles within existing *Physalis*/*Witheringia* TGLs as opposed to increasing the number of TGLs recovered. In

fact, average pairwise amino acid distances for Old World *Lycium* alleles (0.31) compared to New World alleles (0.53) confirms the close relationship of *S-RNase* alleles in the Old World. One allele pair, *L. cinereum* allele CINE4 and *L. oxycarpum* allele OXYC6, have only two synonymous nucleotide differences, making their partial amino acid sequence identical. In ten instances, Old World allele pairs have three or fewer amino acid differences from one another. Matton et al. (1997) has demonstrated that sequences with as few as four amino acid differences can confer different allele specificities. The extent to which the partial alleles reported here are in fact unique alleles remains to be tested, but if so then these data represent an important contribution to our understanding of allele specificity. While many of these Old World individuals have highly similar partial alleles, allele specificities may still be different from one another. Additionally, since only a portion (61%) of the *S-RNase* gene was sequenced in these individuals, other sites under positive selection may be found in parts of the gene that were not sequenced. Igic et. al (2007) found sites under positive selection outside of the regions sequenced here.

Two Old World *Lycium* TGLs, TGLs two and three, each consist of only one *S-RNase* allele. Two *Witheringia*/*Physalis* TGLs had 19 and 38 alleles each, but the third TGL contained a low number (four) of *S-RNase* alleles (Stone and Pierce 2005). One explanation of low numbers of alleles in TGLs could be that heterozygosity at the *S*-locus shelters deleterious mutations to linked genes (Uyenoyama 1997; Richman et al. 2000). This means that deleterious mutations are not expressed at the heterozygous S-locus because there are two different alleles, one that masks the deleterious mutations of the second allele. When individuals containing alleles of recent ancestry (and therefore

containing the same deleterious mutations) mate, this deleterious allelic genetic load is expressed. This may inhibit allele re-diversification in TGLs (Uyenoyama 1997; Richman et al. 2000). In contrast, the loss of genetic load could lead to increased rates of allelic diversification (Paape et al. in press). For example, if genetic load was purged in TGLs one and four, then this could lead to higher allelic diversity within those TGLs. Alternatively, divergent allele advantage could cause the selective retention of a small number of alleles in a TGL (Richman 2000). It may be that it is more difficult for closely related alleles to discriminate self from non-self, therefore more highly divergent alleles are retained in a population to avoid the false rejection of non-self pollen (Uyenoyama 1997; Richman 2000). Chapter one established that *L. ferocissimum* is self-incompatible. The *S-RNase* genealogy (Fig. 2.4) suggests that five *L. ferocissimum* partial alleles are present in TGL one. Although these partial alleles are closely related (FERO2, FERO3 and FERO5 have seven amino acid differences from one another) and fall into the same TGL, they are likely still functionally different *S-RNase* alleles. This suggests that the genetic load may have been purged from alleles in TGL one.

2.4.2 Selection on and Diversity of the *S-RNase* Gene

Site-specific selection analyses on the *S-RNase* gene suggest that both Old World and New World *Lycium* allele sequences are undergoing positive selection (Table 2.5). Strong positive selection on a large number of amino acid sites would be predicted for *S-RNase* alleles re-diversifying following a bottleneck that reduced allelic diversity. Old World alleles have four to eight times more sites under strong positive selection than New World alleles. Presumably, a small number of founding *Lycium* arrived in the Old World with a finite number of alleles when *Lycium* dispersed from the New World. *S-*

RNase alleles may have experienced rapid diversification as a result of positive selection and strong negative frequency dependent selection to increase the number of possible mates.

A large number of sites under positive selection were found beyond conserved region C4 of both the New World and Old World *Lycium* sequences (Fig. 2.6). This has also been found in other Solanaceae S-RNase sequences (Savage and Miller 2006; Igic et al. 2007; Miller et al. in press). These results may suggest that regions of the *S-RNase* gene outside of hypervariable regions HVa and HVb may be involved in allele recognition and encoding allele specificities (Igic et al. 2007). Allele lineages may rely on different or additional amino acid sites for allele recognition in addition to the HVa and HVb regions (Igic et al. 2007).

2.4.3 Identical Partial *S-RNase* Alleles

Seven sets of alleles (OXYC5/FERO6, HIRS6/FERO3, HANT2/BOSC2/OXYC8, OXYC1/HIRS4, CINE5/HIRS5, OXYC7/CINE1 and FERO1/TENU2) from seven Old World *Lycium* species (*L. bosciifolium, L. cinereum, L. ferocissimum, L. hantamense, L. hirsutum, L. oxycarpum* and *L. tenue*) have identical partial coding regions (Table 2.6). The ca. 380 bp region sequenced represents 61% of the total *S-RNase* gene. Out of 36 Old World *Lycium* partial *S-RNase* alleles, only 24 of them had different nucleotide sequences, indicating a large amount of allele sharing. Three allele sets (HANT2/BOSC2/OXYC8, OXYC1/HIRS4, FERO1/TENU2) were shown to have identical (or nearly so) partial coding regions and introns across six species of African *Lycium*. In fact, 50% of the alleles isolated from five species (*L. ferocissimum*, *L. hantamense*, *L. hirsutum*, *L. oxycarpum* and *L. tenue*) are putatively shared with another

species (Table 2.7). This highlights the large amount of allele sharing that is occurring among different species in southern Africa. All identical partial allele sequences were found in the two largest TGLs, TGLs one and four.

Partial allele sharing seems to occur mainly between clades of southern African *Lycium* species. The relationship among Old World *Lycium* species is shown in Figure 2.5. These species form two clades, clades A and B. There is weak bootstrap support (67; Fig. 2.5) for the sister relationship of *L. oxycarpum* and the members of clade A (Fig. 2.5). Allele sharing occurs mainly between clades A and B, or between members of clade A and *L. oxycarpum*. No allele sharing occurs within clade B. *Lycium oxycarpum*, *L. hirsutum* and *L. ferocissimum* each share at least three of their alleles with another species, perhaps because a large number of alleles were sequenced from these individuals. Thus, a larger number of identical partial alleles may be expected with greater sampling and sequencing of other Old World *Lycium* species.

There are several explanations for the extent of partial *S-RNase* sequence sharing observed among *Lycium* species. Since only 61% of the *S-RNase* gene was sequenced, differences between alleles may lie in regions of the gene that were not sequenced. Selection analyses suggest that positive selection occurs in parts of the gene besides the hypervariable regions, therefore it is likely that allele differences lie outside of the regions sequenced here. In support of this, positive selection has been detected in regions outside of the parts of the *S-RNase* gene that were not sequenced in this study (Ioerger et al. 2007). Although additional amino acid differences may lie outside of the regions sequenced, the discovery of identical introns in different *Lycium* species suggests that the entire *S-RNase* gene may be identical. Introns accumulate mutations neutrally through

time, therefore intron sequences should be diverse in different species. Additionally, identical partial allele sequences are found in different species despite strong positive selection in the hypervariable regions and other regions of Old World *S-RNase* alleles.

The discovery of shared partial alleles within different *Lycium* species may also be a result of recent hybridization. The African species *L. hantamense* is a polyploid taxa possibly of hybrid origin (Venter 2000, 2007). It resembles *L. strandveldense*, a polyploid dioecious species, and *L. amoenum*, a diploid cosexual species (Venter 2007); *L. strandveldense* and *L. amoenum* could be potential parental taxa to *L. hantamense*. If this species is a result of recent hybridization, its identical partial alleles may be indicative of *S-RNase* alleles found in parental taxa. This explanation would not apply to allele sharing that occurs among the putatively non-hybrid diploid southern African species (*L. bosciifolium*, *L. oxycarpum*, *L. cinereum*, *L. hirsutum*, *L. ferocissimum* and *L. tenue*), although introgression could explain it.

Present day gene flow among *Lycium* species may also be responsible for allele sharing. Even a limited amount of gene flow between species can initiate an *S-RNase* allele to spread in a population because rare alleles have a selective advantage (Wright 1939; Castric and Vekemans 2004). If gene flow is responsible for the occurrence of identical partial allele sequences among species, then I would expect these taxa to have overlapping ranges and flowering times. Flowering times for most of the species in this study overlap, but times may be sporadic depending on when seasonal African rains occur (Venter 2000). In many instances species ranges also overlap or are in close proximity to one another, although this is not always the case. *L. bosciifolium*, *L. hantamense* and *L. oxycarpum* share identical partial allele BOSC2/HANT2/OXYC8. *L.*

bosciifolium and *L. hantamense* species ranges overlap in their southern regions, but neither of their ranges overlap with or are near the species range for *L. oxycarpum* (Fig. 2.7). Therefore, gene flow is not a likely explanation for allele sharing between *L. bosciifolium*, *L. oxycarpum* and *L. hantamense*.

Lycium is estimated to have dispersed to the Old World relatively recently and therefore the subsequent radiation of African *Lycium* is also relatively recent. Identical partial *S-RNase* sequences among Old World *Lycium* species may be a result of their recent ancestry. The observation that for a given partial allele pair, both the coding region and the introns are identical (or nearly so) is consistent with the rapid and recent radiation of these species.

Allele similarity is present not only in the Old World *Lycium S-RNase* allele data set (results from the present study), but also among the New World *Lycium*. In nine instances, partial allele sequences had four or fewer amino acid differences between alleles in different species. *Lycium cestroides* alleles LCESc and LCESi (which differ from one another by a single bp) have the same partial amino acid (but not nucleotide) sequence as *L. parishii* allele LPAR24. *Lycium cestroides* allele LCESa has one amino acid difference from LPAR25. *Lycium cestroides* alleles LCESd and LCESg are highly similar to *L. parishii* alleles LPAR17 (two differences) and LPAR26 (three differences), respectively. Although none of these comparisons had identical nucleotide sequences (as demonstrated here for Old World species), it is clear that different species contain highly similar alleles. Additional studies have also documented allele similarity among New World Solanaceae species. Savage and Miller (2006) found that two partial *L. parishii* alleles, S_{08} and S_{13}, were identical to two partial *L. andersonii* alleles, S_{11} and S_{01},

isolated by Richman (2000). *Lycium andersonii* S_{11} and *L. parishii* allele S_{08} are also identical to a partial *L. exsertum* and *L. fremontii* allele (Leffler 2006). These individuals have identical coding regions, although only *L. exsertum* and *L. parishii* have identical introns (*L. fremontii* has a one bp difference in its intron and *L. andersonii*'s intron was not sequenced) (Feliciano unpublished data). The partial *S-RNase* nucleotide sequences of *Physalis longifolia* allele S_{26} and *P. cinerascens* allele S_3 are also identical (Lu 2001).

Partial allele sharing may be caused by recent speciation. Identical partial allele sequences are found only between species within the same continent (i.e. between North American species or between Old World species), but not between species on different continents. New World putative allele sharing occurs between closely related species, such as *L. andersonii, L. fremontii, L. parishii* and *L. exsertum* (Fig. 2.8) as does Old World putative allele sharing. This suggests that identical partial allele sequences may be due to recent speciation. Identical partial alleles in different species seem to be a common occurrence and closer scrutiny of all existing and future *S-RNase* sequences may show additional examples of putative shared alleles.

Recently, Surbanovski et al. (2007) suggested that the evolution of new allele specificities may begin with mutations in the *SFB* gene and then subsequent mutations in the *S-RNase* gene. Surbanovski et al. (2007) isolated *S-RNase* alleles from three different species in the Rosaceae family. *S-RNase* alleles from two species, *Prunus tenella* (allele S_8) and *P. avium* (allele S_1) had identical amino acid sequences, whereas *P. dulcis* allele S_{11} had one amino acid difference from the identical alleles. *SFB* (*S*-haplotype-specific F-box) alleles, which are homologous to Solanaceae *SLF* alleles, were also sequenced from these individuals. Surprisingly, SFB sequences corresponding to alleles S_1, S_8 and

S_{11} were not identical. SFB sequences from alleles S_8 and S_{11} differed by three amino acid residues, S_1 and S_{11} differed by 11 amino acid residues and S_1 and S_8 differed by 12 amino acid residues. The *SFB* alleles were able to recognize and reject their self *S-RNase*. Individuals in the Rosaceae contain two introns in their *S-RNase* sequences. Both *S-RNase* allele introns were also sequenced but were found to have little similarity to one another. This suggests that mutations in the *SFB* gene may occur before mutations in the *S-RNase* gene during allele specificity evolution (Surbanovski et al. 2007). This sort of evolutionary process could be occurring in the Old World *S-RNase* alleles sequenced in this study. Due to the young age of Old World *Lycium*, enough time may not have passed for *S-RNase* sequences to accumulate mutations, yet *SLF* allele sequences could be changing. This could be why identical alleles are found in different species, but would not explain why intron sequences, which evolve neutrally and accumulate random mutations through time, are also shared among species.

2.5 Conclusions

The cause of putative allele sharing can only be hypothesized at present. Sequencing a larger portion of the *S-RNase* gene using rapid amplification of cDNA ends (RACE) (Frohman et al. 1988; Igic et al. 2007) will help us determine whether amino acid differences between putatively shared alleles occur outside the regions of the gene that have been sequenced. Isolating the *Lycium SLF* gene would help elucidate whether this gene is also identical among different species or whether the *SLF* gene is mutating before the *S-RNase* gene as found by Surbanovski et al. (2007). Estimating the age of species that share putative alleles can help us determine if identical partial alleles in different taxa are a result of recent ancestry. Determining the rate of evolution of *S-RNase* sequences

and modeling the evolution of sequences will help us discover whether recent ancestry is the cause of shared alleles or if some other ecological process, such as gene flow, explains the existence of putative allele sharing.

Table 2.1. Allele specific primers designed to amplify similar *S-RNase* sequences among African *Lycium* species. All primers are allele specific except for primer ZA-SA-f, which is degenerate and amplified three alleles.

Primer	Sequence (5'-3')
ZA-1f	C G G C A A T G C T A A A T A A C T G C G T C
ZA-1r	T T C A A T G C A C T T G A G G C T A G G A A A C
ZA-2f	A C A C A C C A C T G A A T A A C T G C G C C
ZA-2r	T T G G A A C C G C T C G A G T A A C T C C C
ZA-3f	A G A C A C T A C T G A A T A A C T G C G C C
ZA-4f	T C C G T A A T G C T G A A T A A C T G C C C G
ZA-4r	T C A A C G C A C T T G A G G C T A G G A A C T
ZA-SA-f	C A A G R A S A C A C N A C T R A A T W A C T G C G
ZA-SA-r	C A T T C T T G G G T T A T T T G G A T C A G G
OXYC8u-f	C T C G G C A A T G C T A A A T A A C T G C
OXYC8u-r	C A C A T T T C C A G G A G G T T T T T C

Table 2.2. Allele specific primer combinations used to isolate alleles from genomic DNA from selected *Lycium* individuals.

Individual	Allele(s)	Primers
ZA-21, ZA-23, ZA-38	HANT2/BOSC2/OXYC8	ZA1-f/ZA1-r
ZA-21	BOSC3	ZA4-f/ZA4-r
ZA-23	HANT1	ZA3-f/ZA2-r
ZA-34	TENU1	ZA3-f/ZA2-r
ZA-28, ZA-34	FERO1/TENU2	ZA1-f/ZA1-r
ZA-28	FERO2	ZA2-f/ZA2-r
ZA-38, Genomic *L. hirsutum*	OXYC1/HIRS4/HIRS6	ZA-SA-f/ZA-SA-r

Table 2.3. *S-RNase* alleles, along with their Gen-bank ID numbers, that are included in an alignment of Old World and New World *Lycium*, Solanaceae, and *Antirrhinum S-RNase* sequences. Alleles isolated during this study are indicated in the Gen-bank ID column.

New World *Lycium* Species	Allele	Gen-bank ID
L. andersonii	LAND1	AF105343
L. andersonii	LAND2	AF105344
L. andersonii	LAND5	AF105347
L. andersonii	LAND6	AF105348
L. andersonii	LAND7	AF105349
L. andersonii	LAND11	AF105353
L. andersonii	LAND13	AF105355
L. andersonii	LAND16	AF105358
L. andersonii	LAND17	AF105359
L. andersonii	LAND20	AF105362
L. andersonii	LAND21	AF105363
L. cestroides	LCESa	EU074797
L. cestroides	LCESb	EU074798
L. cestroides	LCESc	EU074799
L. cestroides	LCESd	EU074800
L. cestroides	LCESe	EU074801
L. cestroides	LCESf	EU074802
L. cestroides	LCESg	this study
L. cestroides	LCESh	this study
L. cestroides	LCESi	this study
L. parishii	LPAR1	DQ367853
L. parishii	LPAR2	DQ367854
L. parishii	LPAR3	DQ367855
L. parishii	LPAR4	DQ367856
L. parishii	LPAR5	DQ367857
L. parishii	LPAR6	DQ367858
L. parishii	LPAR7	DQ367859
L. parishii	LPAR8	DQ367860
L. parishii	LPAR9	DQ367861
L. parishii	LPAR10	DQ367862
L. parishii	LPAR11	DQ367863
L. parishii	LPAR12	DQ367864
L. parishii	LPAR13	DQ367865
L. parishii	LPAR14	DQ36786
L. parishii	LPAR15	DQ367867
L. parishii	LPAR17	DQ367868
L. parishii	LPAR18	DQ367869
L. parishii	LPAR19	DQ367870
L. parishii	LPAR21	DQ367871
L. parishii	LPAR22	DQ367872
L. parishii	LPAR23	DQ367873
L. parishii	LPAR24	DQ367874
L. parishii	LPAR25	DQ367875
L. parishii	LPAR26	DQ367876

Solanaceae (*Petunia*, *Solanum* and *Witheringia*) Species	Allele	Gen-bank ID
P. axillaris	PAXI1	AF239908
P. axillaris	PAXI13	AF239909
P. axillaris	PAXI15	AF239910
P. axillaris	PAXI17	AY180050
P. axillaris	PAXInonS	AF239907
P. axillaris	PAXIsc1	AY180048
P. axillaris	PAXIsc2	AY180049
P. cinerascens	PCIN1	AF058930
P. cinerascens	PCIN7	AF058936
P. cinerascens	PCIN11	AF058940
P. crassifolia	PCRA1	L46653
P. crassifolia	PCRA21	L46673
P. integrifolia	PINT2	AF301533
P. integrifolia	PINT6	AF301167
P. integrifolia	PINT7	AF301168
P. integrifolia	PINT8	AF301169
P. integrifolia	PINT9	AF301170
P. integrifolia	PINT10	AF301171
P. integrifolia	PINT11	AF301172
P. integrifolia	PINT12	AF301173
P. integrifolia	PINT13	AF301174
P. integrifolia	PINT15	AF301175
P. integrifolia	PINT16	AF301176
P. integrifolia	PINT17	AF301177
P. integrifolia	PINT19	AF301178
P. integrifolia	PINT20	AF301179
P. integrifolia	PINT21	AF301180
S. carolinense	SCARa	L40539
S. carolinense	SCARb	L40540
S. carolinense	SCARc	L40541
S. carolinense	SCARd	L40542
S. carolinense	SCARe	L40543
S. carolinense	SCARf	L40544
S. carolinense	SCARg	L40545
S. carolinense	SCARh	L40546
S. carolinense	SCARj	L40547
S. carolinense	SCARk	L40548
S. carolinense	SCARn	L40551
S. chacoense	SCHA2	X56896
S. chacoense	SCHA3	X56897
S. chacoense	SCHA11	L36464
S. chacoense	SCHA12	AF176533
S. chacoense	SCHA13	L36667
S. peruvianum	SPERU6	Z26583
W. maculata	WMAC1	AF102071
W. maculata	WMAC2	AF102066
W. maculata	WMAC6	AF102073
W. solanacea	WSOL1	AY454099
W. solanacea	WSOL7	AY454105
W. solanacea	WSOL15	AY454113

Old World *Lycium* Species	Allele	Gen-bank ID
L. afrum	AFRU1	this study
L. bosciifolium	BOSC1	EU074803
L. bosciifolium	BOSC2	EU074804
L. bosciifolium	BOSC3	EU074805
L. bosciifolium	BOSC4	EU074806
L. cinereum	CINE1	EU074807
L. cinereum	CINE2	EU074808
L. cinereum	CINE3	EU074809
L. cinereum	CINE4	EU074810
L. cinereum	CINE5	EU074811
L. cinereum	CINE6	EU074812
L. ferocissimum	FERO1	EU074813
L. ferocissimum	FERO2	EU074814
L. ferocissimum	FERO3	EU074815
L. ferocissimum	FERO4	EU074816
L. ferocissimum	FERO5	EU074817
L. ferocissimum	FERO6	this study
L. gariepense	GARI1	this study
L. hantamense	HANT1	this study
L. hantamense	HANT2	this study
L. hirsutum	HIRS1	EU074818
L. hirsutum	HIRS2	EU074819
L. hirsutum	HIRS3	EU074820
L. hirsutum	HIRS4	this study
L. hirsutum	HIRS5	this study
L. hirsutum	HIRS6	this study
L. oxycarpum	OXYC1	EU074821
L. oxycarpum	OXYC2	EU074822
L. oxycarpum	OXYC3	EU074823
L. oxycarpum	OXYC4	EU074824
L. oxycarpum	OXYC5	EU074825
L. oxycarpum	OXYC6	EU074826
L. oxycarpum	OXYC7	this study
L. oxycarpum	OXYC8	this study
L. tenue	TENU1	this study
L. tenue	TENU2	this study

Outgroup (*Antirrhinum*)	Allele	Gen-bank ID
A. hispanicum	AHISP2	X96465
A. hispanicum	AHISP4	X96466
A. hispanicum	AHISP5	X96464

Table 2.4. Genotypes of Old World *Lycium.* Genotypes and individual collection numbers for individuals isolated from this study and from Miller et al. (in press) are shown. Shared sequences are indicated with forward slashes and are enclosed by parentheses (i.e. identical partial alleles HIRS4 found in *L. hirsutum* and OXYC1 found in *L. oxycarpum* appear as HIRS4/OXYC1). Partial alleles isolated from this study are highlighted in gray.

Species	Collection Number	Genotype
Lycium afrum	ZA32	AFRU1
Lycium bosciifolium	ZA21	(BOSC2/HANT2/OXYC8) and BOSC3
Lycium bosciifolium	ZA22	BOSC1 and BOSC4
Lycium cinereum	ZA07	(CINE1/OXYC7) and CINE2
Lycium cinereum	ZA08	CINE6
Lycium cinereum	ZA09	(CINE1/OXYC7) and (CINE5/HIRS5)
Lycium cinereum	ZA35	CINE3 and CINE4
Lycium ferocissimum	WCNP01	(FERO3/HIRS6) and FERO4
Lycium ferocissimum	ZA29	(FERO6/OXYC5)
Lycium ferocissimum	ZA30	FERO2 and FERO5
Lycium ferocissimum	ZA31	(FERO3/HIRS6) and FERO4
Lycium ferocissimum	ZA28	(FERO1/TENU2) and FERO2
Lycium gariepense	ZA20	GARI1
Lycium hantamense	ZA23	HANT1 and (HANT2/BOSC2/OXYC8)
Lycium hirsutum	ZA01	HIRS3
Lycium hirsutum	ZA03	HIRS1 and HIRS2
Lycium hirsutum	ZA02	(HIRS6/FERO3) and (HIRS4/OXYC1)
Lycium hirsutum	ZA04	(HIRS5/CINE5) and (HIRS6/FERO3)
Lycium hirsutum	ZA05	(HIRS4/OXYC1) and (HIRS6/FERO3)
Lycium oxycarpum	ZA36	(OXYC1/HIRS4) and (OXYC5/FERO6)
Lycium oxycarpum	ZA39	OXYC4 and OXYC6
Lycium oxycarpum	ZA41	OXYC3 and OXYC2
Lycium oxycarpum	ZA43	OXYC2 and OXYC3
Lycium oxycarpum	ZA37	(OXYC1/HIRS4)
Lycium oxycarpum	ZA38	(OXYC1/HIRS4) and (OXYC8/HANT2/BOSC2)
Lycium oxycarpum	ZA42	(OXYC7/CINE1)
Lycium tenue	ZA34	(TENU2/FERO1) and TENU1

Table 2.5. Likelihood ratio test results, parameter values and numbers of positively selected sites for four models of selection [M1a-nearly neutral, M2a-positive, M7(beta)-nearly neutral, M8(beta&ω>1)-positive] in New World and Old World *Lycium*. For models M1 and M2a, the inferred proportions of sites evolving are indicated by parameters p_0 (purifying selection) and p_1 (neutral evolution). For model M2a, the parameter p_2 (positive selection) is also given. The d_N/d_S rate ratios for each proportion of sites evolving for models M1 and M2a are ω_0 (purifying selection) and ω_1 (neutral selection). The d_N/d_S rate ratio ω_2 (positive selection) is also given for model M2a. Parameters p and q for beta distribution models M7 and M8 describe the d_N/d_S rate ratio distribution between 0 and 1. In model M8, the parameter p_0 is the proportion of sites within the beta distribution, p_1 is the proportion of positively selected sites and ω_s is the d_N/d_S rate ratio for the positively selected sites.

Lycium	Model	ℓ	$2\Delta\ell$ (df, P-value)	Parameters	No. of positively selected sites
Old World	M1a: nearly neutral	-2443.7112	46.412 (2, < 0.0000001)	p_0=0.378 (ω_0=0.106), p_1=0.622 (ω_1=1.00)	N/A
	M2a: positive	-2420.5054		p_0=0.302 (ω_0=0.107), p_1=0.512 (ω_1=1.00), p_2=0.187 (ω_2=4.68)	4 (1.0 ≥ P ≥ 0.99) 4 (0.99 > P ≥ 0.95) 2 (0.95 > P ≥ 0.90) 11 (0.90 > P ≥ 0.50) 21 Total
	M7 (beta): nearly neutral	-2447.4323	50.154 (2, < 0.0000001)	p=0.339, q=0.201	N/A
	M8 (beta&ω>1): positive	-2422.3554		p_0=0.792, p_1=0.207 (ω_s=4.36); p=0.341, q=0.197	5 (1.0 ≥ P ≥ 0.99) 6 (0.99 > P ≥ 0.95) 4 (0.95 > P ≥ 0.90) 15 (0.90 > P ≥ 0.50) 30 Total
New World	M1a: nearly neutral	-7667.5517	14.182 (2, < 0.0008326)	p_0=0.477 (ω_0=0.209), p_1=0.523 (ω_1=1.00)	N/A
	M2a: positive	-7660.4608		p_0=0.460 (ω_0=0.220), p_1=0.445 (ω_1=1.00), p_2=0.095 (ω_2=1.89)	1 (0.99 > P ≥ 0.95) 2 (0.95 > P ≥ 0.90) 6 (0.90 > P ≥ 0.50) 9 Total
	M7 (beta): nearly neutral	-7647.2755	16.537 (2, < 0.0002565)	p=0.592, q=0.494	N/A
	M8 (beta&ω>1): positive	-7639.007		p_0=0.819, p_1=0.181 (ω_s=1.45); p=0.708, q=0.809	4 (0.99 > P ≥ 0.95) 2 (0.95 > P ≥ 0.90) 13 (0.90 > P ≥ 0.50) 19 Total

Table 2.6. Alleles found in nine southern African *Lycium* species. Alleles obtained from this study are underlined. All other alleles were isolated in Miller et al. (in press). Identical alleles are highlighted in the same color. Partial alleles with the superscript "i" have identical coding and intron sequences. Partial alleles with an asterisk, (i.e. i*) differ between introns by 1 bp. Material was not available to isolate partial sequences from genomic DNA in the remaining shared partial allele pairs.

Lycium afrum	*Lycium hantamense*	*Lycium bosciifolium*	*Lycium oxycarpum*	*Lycium cinereum*	*Lycium hirsutum*	*Lycium ferocissimum*	*Lycium tenue*	*Lycium gariepense*
AFRU1	HANT1	BOSC1	**OXYC7**	**CINE1**	HIRS1	**FERO1^{i*}**	**TENU2^{i*}**	GARI1
	HANT2^{i}	**BOSC2^{i}**	**OXYC8^{i}**	CINE2	HIRS2	FERO2	TENU1	
		BOSC3	OXYC2	CINE3	**HIRS6**	**FERO3**		
		BOSC4	**OXYC1^{i}**	CINE4	**HIRS4^{i}**	FERO4		
			OXYC5	CINE6	HIRS3	**FERO6**		
			OXYC3	**CINE5**	**HIRS5**	FERO5		
			OXYC4					
			OXYC6					

Table 2.7. The number of individuals, alleles and putatively shared alleles isolated from nine southern African *Lycium* species. The percentage of putatively shared alleles for each *Lycium* species is indicated.

Lycium Species	Number of Individuals	Number of Alleles	Number of Alleles Shared	Percent of Alleles Shared
L. afrum	1	1	0	0
L. bosciifolium	2	4	1	25
L. cinereum	4	6	2	33
L. ferocissimum	5	6	3	50
L. gariepense	1	1	0	0
L. hantamense	1	2	1	50
L. hirsutum	5	6	3	50
L. oxycarpum	7	8	4	50
L. tenue	1	2	1	50
Totals:	27	36	15	

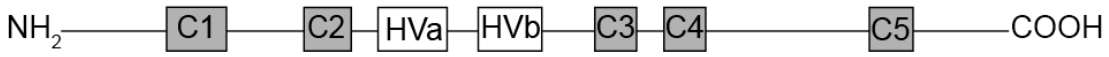

Figure 2.1. Diagram of the *S-RNase* gene in Solanaceae. Conserved Regions (C1-C5, orange) and hypervariable (HVa and HVb, blue) regions are indicated. The *S-RNase* gene is ca. 620 bp long in a multi species alignment of Solanaceae (Ioerger 1991) and contains a single intron that is nested within hypervariable region HVa.

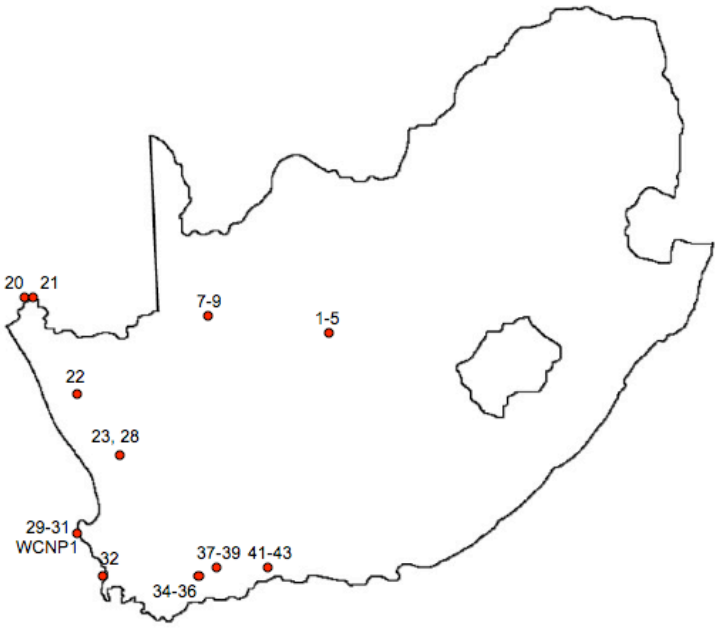

Figure 2.2. Southern African *Lycium* accessions used in this study. Red dots indicate the locations of plants that style collections were made from in 2005 and 2006. Collection numbers are given without the ZA prefix (i.e. ZA-22 = 22).

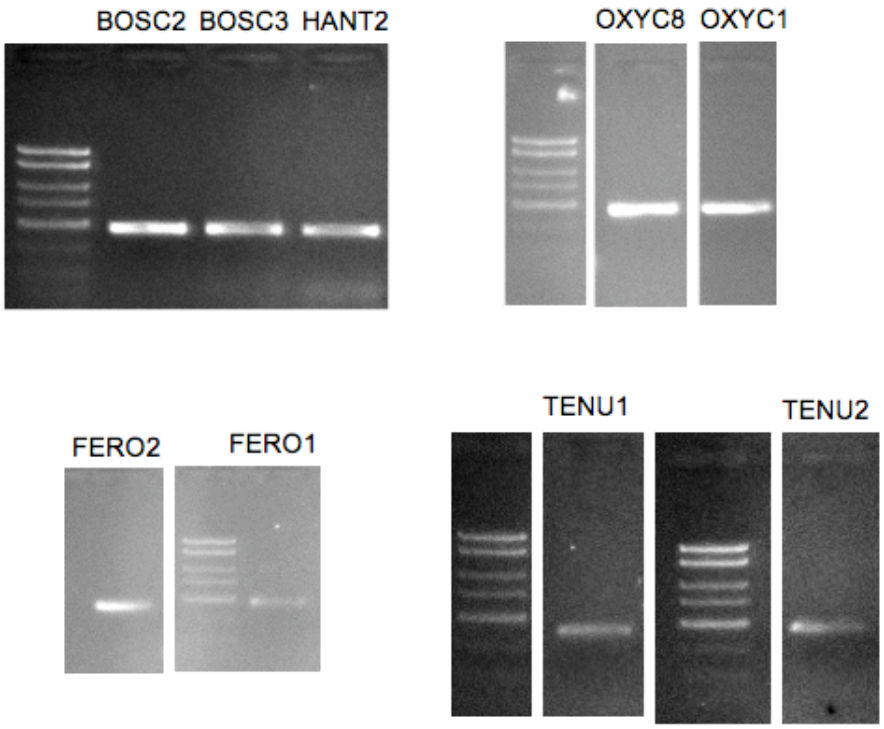

Figure 2.3. Positive amplification of identical partial alleles from genomic DNA. Identical partial alleles, including the intron, were isolated from genomic DNA to isolate three sets of putatively shared alleles. Gel pictures for the positive amplification of two of the allele sets (BOSC2/HANT2/OXYC8 and FERO1/TENU2) are shown. The second allele in the *L. bosciifolium* individual that contained allele BOSC2, the *L. oxycarpum* individual that contained allele OXYC8, the *L. ferocissimum* individual that contained allele FERO1 and the *L. tenue* individual that contained allele TENU2 were also amplified. The fifth band down on each gel ladder picture is 500 bp long.

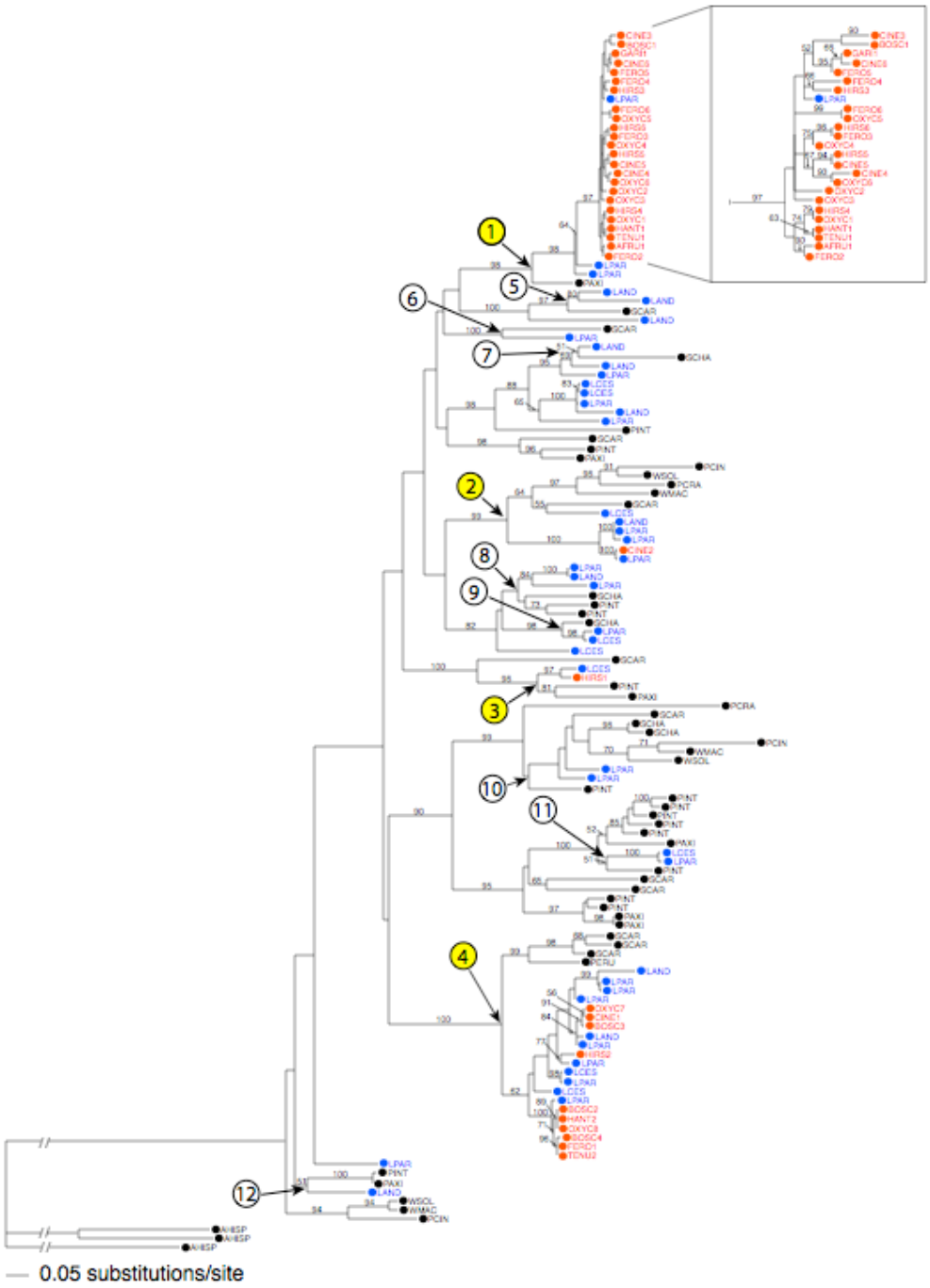

Figure 2.4. Maximum likelihood S-RNase genealogy constructed using New and Old World *Lycium* and Solanaceae *S-RNase* alleles. Maximum likelihood bootstrap values are indicated. Red dots indicate Old World *Lycium S-RNase* and blue dots indicate New World *Lycium* alleles. All other Solanaceae alleles are indictaed by a black dot. (See Table 2.3 for allele identifications.) Trans-generic lineages (TGLs) are numbered one through 12. TGLs that include both Old and New World *Lycium* are highlighted in yellow. Slash marks indicate where branches were shortened. Old World and New World *Lycium* alleles in TGL one are amplified in an inset to show allele relationships.

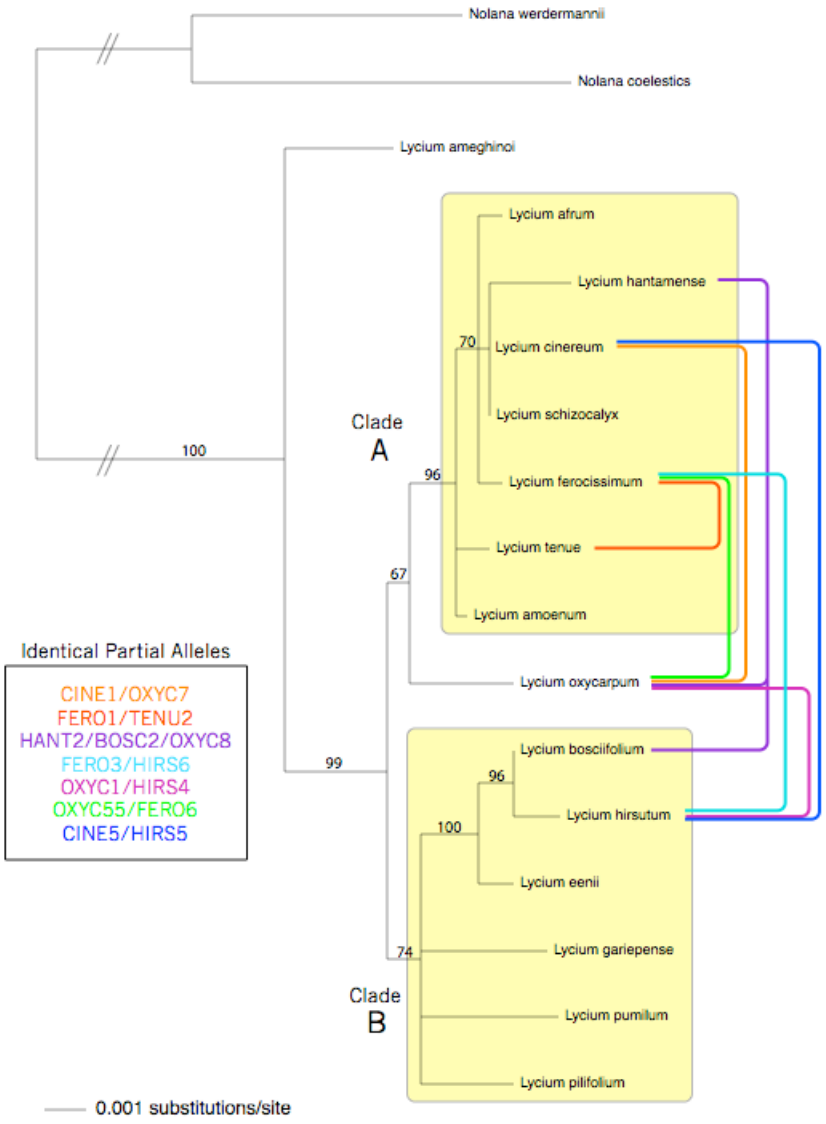

Figure 2.5. *Lycium* ML phylogeny using nuclear and chloroplast genes. Nuclear gene GBSSI and the chloroplast genes *trnD-trnT* and *trnT-trnL* were used to create a maximum likelihood phylogeny. ML bootstrap values for each node are indicated above branches. Slash marks indicate where branch lengths were shortened for presentation. Outgroup species are *Nolana werdermanii* and *N. coelestics*. *Lycium ameghinoi* is the only New World *Lycium* species; all remaining *Lycium* are from the Old World. Yellow shaded boxes surround African *Lycium* clades A and B. Lines indicate putative sharing of identical partial S-RNase allele sequences among species. Individuals connected by the same colored line share the same allele. Identical partial alleles in the legend are the same color as the line connecting the species they are found in.

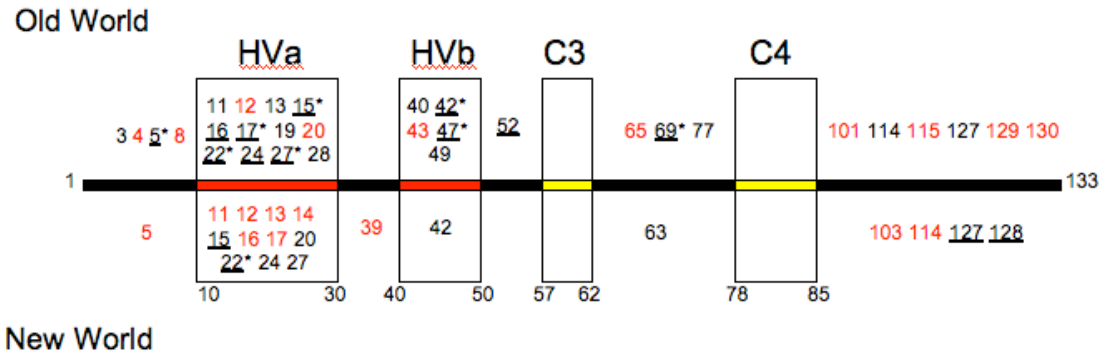

Figure 2.6. Positive selection on the *S-RNase* gene. The amino acid positions of sites under positive selection are indicated above and below the *S-RNase* gene. The partial *S-RNase* gene amino acid positions range from one to 133. Hypervariable regions HVa (positions 10-30) and HVb (positions 40-50) and conserved regions C3 (positions 57-62) and C4 (positions 78-85) are indicated. Positions above the gene are positively selected sites in Old World *Lycium* and below the gene are positively selected sites in New World *Lycium*. Numbers in black are inferred to be under positive selection according to the PAML *codeml* M2a and M8 (beta&ω) models of positive selection. Numbers in red are sites that are inferred to be under positive selection for only model M8. Sites with an asterisk have a greater than 95% probability of being under positive selection for the M2a model. Sites that are underlined have a greater than 95% probability that the site is under positive selection for the M8 model.

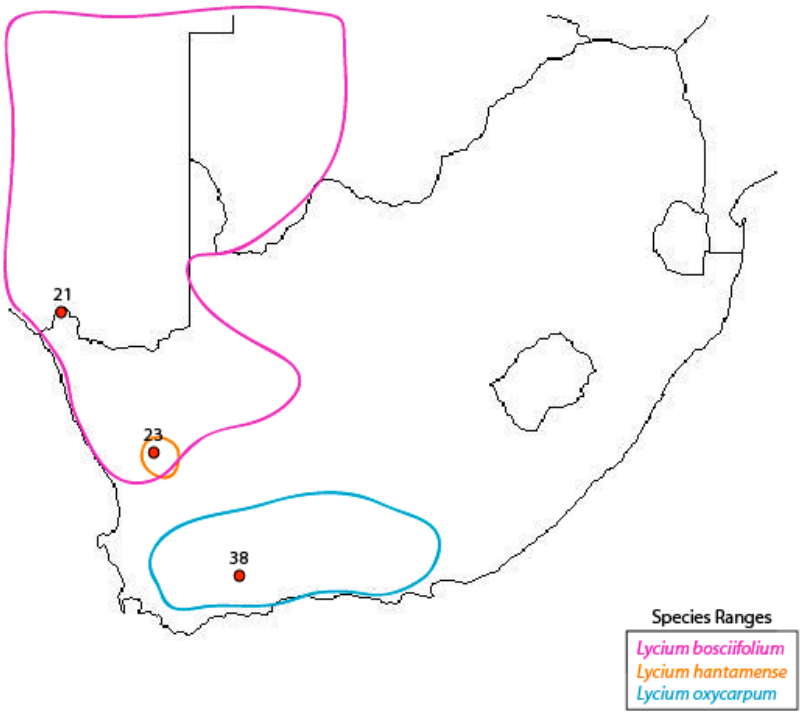

Figure 2.7. Species ranges for *L. bosciifolium*, *L. hantamense* and *L. oxycarpum* in Southern Africa. Locations of individuals that have identical partial allele sequences are indicated (*L. bosciifolium* = 21, *L. hantamense* = 23, *L. oxycarpum* = 38). *L. bosciifolium's* species range is in pink, *L. hantamense's* species range is in orange and *L. oxycarpum's* species range is in turquoise.

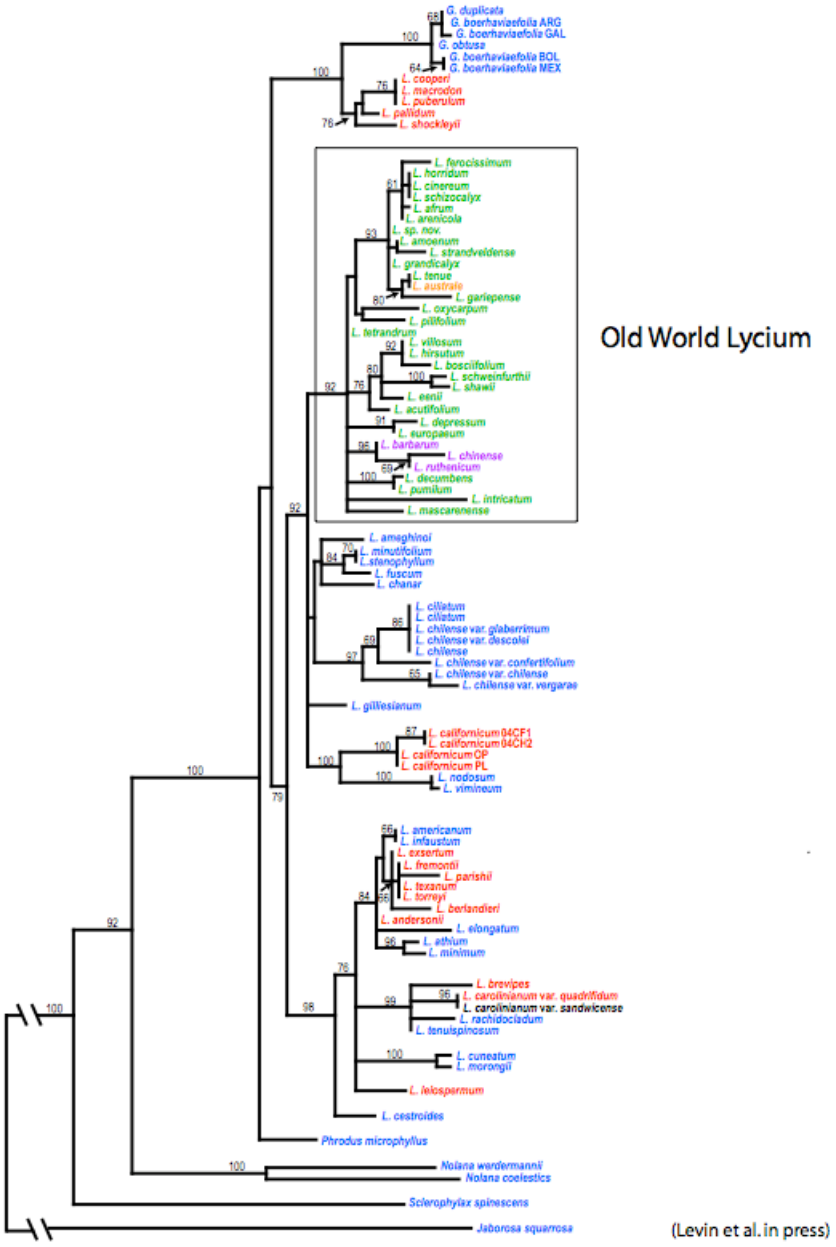

Figure 2.8. Maximum likelihood phylogeny for *Lycium* using nuclear GBSSI sequences (adopted from Levin et al. 2007). Species are color coded according to the continent they reside in. Red=North America, Blue=South America, Green=Africa, Orange=Australia, Purple=Asia, Black=Pacific Islands. A box is drawn around Old World *Lycium* species. Old World *Lycium* form a highly supported monophyletic clade.

CHAPTER 3

ISOLATION OF THE *LYCIUM SLF* GENE

3.1 Introduction

The *SLF* (*S*-locus F-box) gene, which is expressed in pollen grains of gametophytic self-incompatible (GSI) plants, is the pollen component of a set of two genes that interact with each other in pollen tubes. The *SLF* and *S-RNase* (expressed in style tissue) genes are involved in self-recognition of the GSI response. When allele specificities of both the *SLF* and *S-RNase* gene are the same, the self-incompatibility reaction is triggered, leading to the termination of pollen tube growth. The *S-RNase* and *SLF* genes are tightly linked at the *S*-locus (Murfett et al. 1994) and in *Petunia hybrida* are found in a centromeric location, which inhibits recombination between the two genes (Entani et al. 1999). In the Solanaceae and Plantaginaceae families, the pollen *S*-gene has been named *SLF* (Lai et al. 2002) whereas in the Rosaceae it has been named *SFB* (*S*-haplotype-specific F-box; Entani et al. 2003). The role *SLF* plays in the abortion of pollen tube growth and how *SLF* specificities interact with S-RNases is an area of active investigation (McClure 2006).

The *SLF*/*SFB* genes have only recently been discovered. The *SLF* gene was isolated in *Petunia* (Solanaceae) by sequencing a 328 kb region at the *S*-locus around the *S-RNase* gene (Wang et al. 2003; Wang et al. 2004). This was also done in *Prunus dulcis* (Rosaceae; 72 kb sequenced), *P. mume* (Rosaceae; 65 kb sequenced) (Ushijima et al. 2001; Entani et al. 2003) and *Antirrhinum* (Plantaginaceae; 65 kb sequenced) in search of the pollen *S*-gene (Lai et al. 2002). Researchers found that polymorphic F-box genes,

named *SLF* and *SFB,* were the most-likely candidates for the pollen *S*-gene (Wang et al. 2003; Wang et al. 2004; Lai et al. 2002; Ushijima et al. 2003). The first pollen *SLF* gene reported was in *Antirrhinum* (Lai et al. 2002), but was found later in *Prunus* (Entani et al. 2003; Ushijima et al. 2003) and *Petunia* (Wang et al. 2003; Wang et al. 2004). *SLF* genes are composed of conserved and hypervariable regions, much like the *S-RNase* gene, suggesting that *SLF* plays a role in allele recognition (Ioerger et al. 1991; Matton et al. 1997; Ikeda et al. 2004; Vaughan et al. 2006). The *SLF* gene was shown to be expressed in pollen and interact with the *S-RNase* gene in an allele-specific manner in *Petunia* (Sijacic et al. 2004) and *Antirrhinum* (Qiao et al. 2004b), definitively defining *SLF*/*SFB* as the pollen *S*-gene.

An F-box motif is the main characterizing feature of the *SLF* gene. This F-box is found in many different genes and is a site of protein-protein interaction (Kipreos and Pagano 2000). F-box proteins are typically part of SCF (Skp I, Cullin and F-box protein) ubiquitin-ligase complexes and bind substrates for ubiquitin-mediated protein degradation (Kipreos and Pagano 2000). In vivo, the SCF protein complex and F-box protein are hypothetically joined together by the F-box motif (Kipreos and Pagano, 2000); specific proteins can be targeted by the F-box (McClure 2004). The F-box motif is found in the *SLF* gene's N-terminal region (Lai et al. 2002). Regions outside the F-box domain are where substrate specificity occurs (Lai et al. 2002). Allelic variants of the pollen *SLF* gene are present in GSI populations, much like *S-RNase* alleles (Lai et al. 2002), and these variants interact with S-RNases (Qiao et al. 2004). *SLF* polymorphism is similar to that of *S-RNase* in *Prunus* (Sassa et al. 2007) and the most variable regions of the *SLF* gene are found near its carboxyl-terminus (Ikeda et al. 2004). PCR analysis

shows that the *SLF* gene maps to the *S*-locus (Lai et al. 2002), and in *P. inflata* is approximately 1.3 kb long and shows 10.3-11.6% sequence variability (Sijacic et al. 2004).

Several non-*S*-specific factors that play a role in gametophytic self-incompatibility have also been identified. The *HT-B* protein is required for self-incompatibility in Solanaceae (Goldraij et al. 2006; Cruz-Garcia et al. 2003) and a polymorphic arabinogalactan protein (AGP) named *120k* is also necessary for *S*-specific pollen rejection (Hancock et al. 2005; McClure 2006). *HT-B* levels decline significantly in pollen tubes after compatible pollinations (Goldraij et al. 2006) and pollen rejection fails when 120k is absent (Hancock et al. 2005). However, how these proteins are involved in the self-incompatibility response is not fully understood (reviewed in McClure 2006).

In the Solanaceae, the *SLF* gene has been isolated in only one genus, *Petunia* (Sijacic et al. 2004; Wang et al. 2004; Tsukamoto et al. 2005). *Petunia* is located within the subfamily Petunioideae, which is basal to the *Lycium* subfamily Solanoideae (Olmstead et al. 1999). In the present study, the *SLF* gene was isolated from *Petunia inflata* following the protocol of Tsukamoto et al. (2005) and attempts were made to isolate *SLF* from *Lycium* (Solanaceae) using an allele-specific and degenerate primer approach. Several studies have used primers designed from different species in the same genera to isolate the *SLF* gene (Tsukamoto et al. 2005; Vaughan et al. 2006; Surbanovski et al. 2007). Tsukamoto et al. (2005) designed primers from *P. inflata* to isolate *SLF* from *P. axillaris* and Vaughan et al. (2006) designed primers from *Prunus dulcis* and *P. mume* (Rosaceae) to isolate the *SFB* gene from *P. avium*. *S-RNase* alleles have shared ancestral polymorphism, which means that an *S-RNase* allele in one genus may be more similar to

an *S-RNase* allele in a different genus than to a congeneric allele (Ioerger et al. 1990; Richman et al. 1995; Igic et al. 2001). Because of this trans-generic polymorphism, some *Lycium S-RNase* alleles are highly similar to alleles found in *Petunia*. This means *Lycium SLF* alleles may also be very similar to *Petunia SLF* alleles. Allele-specific and degenerate primers were designed from *P. inflata* and *P. axillaris SLF* alleles, as well as *Solanum lycopersicon* and *S. tuberosum* (both in the Solanaceae family) putative *SLF* alleles, in an attempt to isolate *Lycium SLF*.

3.2 Methods

3.2.1 Study System

Lycium is a member of the Solanaceae family. Members of this genus are perennial shrubs found throughout the world. *Lycium* live in dry, saline environments on every continent but Antarctica. Three areas of the world, the North American southwest, western South America and southern Africa, contain the majority of the world's *Lycium* species. About 30 species of *Lycium* are found in South America (Bernardello 1986; Hitchcock 1932; Levin and Miller 2005), 21 species in North America (Hitchcock 1932; Chiang-Cabrera 1981; Miller and Venable 2000, 2003; Miller 2002) and 26 species in southern Africa (Venter 2000, 2007). Old World *Lycium* form a monophyletic clade within a group of American *Lycium* species; North and South American *Lycium* are not monophyletic (Levin and Miller 2005; Levin et al. 2007). This suggests that there was a single dispersal of *Lycium* to the Old World (Levin and Miller 2005; Levin et al. 2007; Miller et al. 2008).

3.2.2 *SLF* Isolation

Undehisced anthers were collected from flowers of *Lycium fremontii* and *L. andersonii* plants in Arizona during March 2006 and immediately preserved in RNAlater® (Ambion, Inc., Austin, TX). Anthers from one *L. fremontii* (sample number 21) and four *L. andersonii* (sample numbers 9, 48, 67 and 80) plants were used in the present study. Because the *SLF* gene has been isolated from *Petunia*, *Petunia inflata* was used as a positive control during my attempts at isolating the *Lycium SLF* gene. *Petunia inflata* seeds were obtained from Nijmegen Botanical Garden (accession number 954750061) and plants were grown in the Amherst College greenhouse. *Petunia* anthers were collected during June 2006 and immediately preserved in RNAlater.

Total RNA was extracted from an average of 18 anthers using a Qiagen RNeasy Plant Mini Kit (Qiagen, Inc., Valencia, CA) and cDNA was synthesized from RNA using a Novagen First Strand cDNA Synthesis Kit (EMD Biosciences, Inc., Madison, WI). Forward primer SLF-1f (5'- GATCCAGTTGAATGAGATGGA -3'; taken from Tsukamoto et al. 2005) and reverse primer SLF-8r (5'- CAATCAAACCATCACAAGGAC -3'; designed from conserved regions of the *Petunia inflata SLF* gene) were used to amplify *SLF* alleles in *P. inflata*. For a 50 µl PCR reaction, 5 µl cDNA, 1 µl of 10mM dNTPs, 2.5 µl of 10µM primers, 0.40 µl Taq DNA Polymerase and 5 µl 10X Taq buffer were used. PCR conditions were as follows: 1 cycle at 94°C (2 min.); 30 cycles at 92°C (30 sec.), 45°C (40 sec.), 72°C (1:30 min.); 1 cycle at 72°C (10 min.).

Allele specific and degenerate primers were designed from conserved regions of known *Petunia inflata* and *P. axillaris*, and putative *Solanum lycopersicon* and *S.*

tuberosum SLF alleles or were taken from the literature (Tsukamoto et al. 2005; Tables 3.1 and 3.2). Sixteen primer pairs, consisting of eight allele-specific (SLF-1f, SLF-2f, SLF-3f, SLF-5f, SLF-6r, SLF-7r, SLF-8r, SGN-f and SGN-r) and eight degenerate (Sd-f, Sd-r, Sd-fi, Sd-ri, SLF-DEGNL1f, SLF-DEGNL2r, SLF-DEG3f and SLF-DEG6r) primers were used in amplifications of cDNA synthesized following total RNA extraction of anthers (Table 3.3). The concentration and amounts of PCR reagents and the PCR amplification cycles used for each primer combination are in Table 3.4.

Amplification products were cloned using the Novagen pT7Blue Perfectly Blunt Cloning Kit (EMD Biosciences, Inc., Madison, WI) to separate putative *SLF* alleles. White colonies and colonies with a faint blue dot were lifted with sterile pipette tips and suspended in 50 µl of water. Colonies were boiled at 99°C for five minutes and then centrifuged at 13,000 RPM for one minute. The supernatent was then transferred into a new microcentrifuge tube.

Colony DNA was screened for putative *SLF* alleles by amplifying 5 µl of DNA in a 25 µl PCR reaction (see Table 3.4 for PCR reagent concentrations and conditions). Colony DNA containing a putative *SLF* allele was amplified with vector specific primers U-19 (5'-GTT TTC CCA GTC ACG ACG T-3') and R-20 (5'-CAG CTA TGA CCA TGA TTA CG-3') with the following PCR conditions: 95°C (5 min.); 2 cycles at 94°C (30 sec.), 55°C (1 min.), 72°C (2 min.); 2 cycles at 94°C (30 sec.), 54°C (1 min.), 72°C (2 min.); 2 cycles at 94°C (30 sec.), 53°C (1 min.), 72°C (2 min.); 30 cycles at 94°C (30 sec.), 52°C (1 min.), 72°C (2 min.); 72°C (5 min.). For a 50µl PCR reaction, 10 µl of cloned genomic, 1 µl of 10 mM dNTPs, 1 µl each of 5 pmol/µl primers R-20 and U-19 and 0.25 µl NEB Taq DNA polymerase were used. DNA was cleaned using a Qiagen

PCR Purification Kit (Qiagen, Inc., Valencia, CA), quantified on a 1% agarose electrophoresis gel using the Bio-Rad E-Z Load Precision Molecular Mass Standard (Bio-Rad laboratories, Inc., Hercules, CA), and sequenced on an Applied Biosystems Automated 3730 DNA Analyzer by the Biotechnology Resource Center at Cornell University (Ithaca, NY). Sequences were aligned by eye in Sequence Alignment Editor (Se-Al) version 2.0a11 (Rambaut 1996).

Gel extraction was used to isolate putative *SLF* alleles in several accessions of *Lycium andersonii* (samples 9, 48, 67 and 80). First, *L. andersonii* cDNA was amplified with allele-specific primers SLF-2f and SLF-8r and PCR amplification products were run on a 1.5% agarose gel. DNA bands of the appropriate size (ca. 450 bp) were cut from the gel using a clean razor blade and DNA was extracted from the gel using a Qiagen QIAquick gel extraction kit (Qiagen, Inc., Valencia, CA). DNA from the gel extraction was then amplified a second time to increase the amount of amplification product.

3.3 Results

The pollen *SLF* gene was successfully amplified from *Petunia inflata*. A 465 bp long PCR amplification product was isolated from *P. inflata* sample 8 and *P inflata* sample 1 pollen using forward primer SLF-1f and reverse primer SLF-8r (Fig. 3.1), but was sequenced from only *P. inflata* sample 8. The allele contained an F-box motif and was extremely similar (7 bp difference) to the PaSLF-19 allele isolated by Tsukamoto et al. (2005).

RNA was successfully isolated from *L. fremontii* individual 21 and *L. andersonii* individuals 9, 48, 67 and 80, but attempts to isolate the *SLF* gene from *Lycium andersonii* and *L. fremontii* were unsuccessful. Sixteen different primer pairs were used to attempt

to amplify the *SLF* gene in *Lycium* (Table 3.3). The majority of PCR amplifications using allele-specific and degenerate primers failed or produced bands that were not the correct *SLF* size (Table 3.5). PCR amplifications using primers SLF-2f and SLF-8r produced amplification products that contained multiple bands, one of which was the size of the putative *SLF* gene. Gel extractions and subsequent PCR amplifications (using SLF-2f and SLF-8r) from *L. andersonii* 9, 48, 67 and 80 were not successful. DNA bands of the appropriate size were isolated from *L. andersonii* 9, 48, 67 and 80 and *L. fremontii* 21 using primer pairs SFL-1f/SLF-8r (*L. fremontii* 21; ca. 460 bp), SD-fi/SD-ri (Fig. 3.2; *L. andersonii* 9, 48, 67 and 80; ca. 340 bp) and SLF-5f/SLF-7r (*L. andersonii* 48, 67 and 80; ca. 650 bp). When the primer sequences were removed from the cloned DNA sequences, these individuals did not show any similarity to known *SLF* alleles when using the nucleotide basic local alignment search tool (BLAST) on the National Center for Biotechnology Information (NCBI) website (Table 3.6).

3.4 Discussion

The *Petunia SLF* gene was isolated using an allele-specific primer approach, but the *Lycium SLF* gene has proven to be elusive using this same technique. Sixteen different primer pair combinations were used in attempt to isolate the *Lycium SLF* gene. Although some positive bands of the correct size were obtained, sequenced *Lycium* DNA did not show homology to known *SLF* genes or F-box sequences.

The *SLF* gene has been isolated from only a limited number of individuals in the Solanaceae. If a large amount of allele diversity exists among *SLF* alleles, as there is among *S-RNase* alleles, it may be very difficult to isolate a variety of alleles using allele-specific primers. Additionally, since the degenerate primers used in this study were

designed from a small number of *SLF* alleles, it may be difficult to isolate a variety of alleles using these primers as well. My PCR approach to isolating the *SLF* gene may have been unsuccessful because the primers used were too allele-specific for use on a limited number of specimens. Therefore, if this PCR approach was used on a greater sampling of *Lycium* individuals, there may be a greater chance of isolating a *Lycium SLF* allele.

The pollen *SFB* and style *S-RNase* genes were recently isolated from *Prunus tenella*, *P. dulcis* and *P. avium* (Surbanovski et al. 2007). *Prunus tenella* (S_8), *P. dulcis* (S_{11}) and *P. avium* (S_1) each contained an *S-RNase* allele that were highly similar to one another. The *P. tenella S-RNase* allele had the same amino acid sequence as the *P. avium* allele and one amino acid difference from the *P. dulcis* allele. *SFB* alleles isolated from these individuals did not show as high of sequence similarity. The *P. tenella* and *P. dulcis* alleles differed by three amino acids, *P. dulcis* and *P. avium* alleles differed by 11 amino acids and *P. tenella* and *P. avium* alleles differed by 12 amino acids. The *Prunus S-RNase* alleles have the same amino acid sequence, yet the *SFB* alleles do not. Vaughan et al. (2006) found that the genealogies of *P. avium S-RNase* and *SFB* alleles obtained from the same individuals are different. This could mean that *SLF* and *SFB* alleles evolve differently than *S-RNase* alleles and relying on the phenomenon of trans-generic polymorphism to isolate *Lycium* alleles using *Petunia* primers may be difficult.

3.5 Conclusions

Whereas I was able to isolate *SLF* from *P. inflata*, I was not successful in isolating the *Lycium SLF* gene, but further work using PCR methods may be successful. If allele-specific and degenerate primers designed from *Petunia* are used on a larger number and

greater diversity (i.e. different species) of *Lycium*, a *Lycium SLF* allele similar to *Petunia* alleles may be found. Conserved and hypervariable regions of the *SLF* gene have been identified in *Prunus* (Sassa et al. 2007), allowing for the opportunity to construct better primers from conserved regions of the gene. Alternatively, I could target a *Lycium* individual that contains an *S-RNase* allele that shares a TGL with a *Petunia* allele. Because the *Lycium* and *Petunia S-RNase* alleles share ancestry, their *SLF* alleles may also be similar due to ancestry. I can create primers based on the *Petunia SLF* allele from the individual that shares a TGL with *Lycium* to try to isolate the *Lycium SLF* allele.

The rapid amplification of cDNA ends (RACE) can also be used to amplify the *SLF* gene. In RACE, a homopolymeric tail (a string of a single type of nucleotide) is tagged with a known nucleotide sequence that is added during cDNA synthesis. One primer that amplifies from a known portion of the *SLF* gene and one primer that amplifies the tagged sequence can be used to amplify the tagged cDNA end in a PCR reaction. If even just a portion of a conserved region of the *SLF* gene can be amplified, RACE can amplify the remaining 5' or 3' portion of the gene.

Alternatively, additional molecular based approaches, such as coimmunoprecipitation assays, can be used to isolate the *Lycium SLF* gene. Sequencing of the region around the *S-RNase* gene has been successful in isolating *SLF* and *SFB* in several studies and would be useful here (Wang et al. 2003; Wang et al. 2004; Ushijima et al. 2001; Entani et al. 2003; Lai et al. 2002). Further research into the structure and function of the *SLF* gene may shed light on alternative ways of isolating *SLF* from different genera.

Table 3.1. Allele-specific and degenerate *SLF* primers. Allele-specific primers are highlighted in gray. Degenerate primers are not highlighted. IUPAC codes for degenerate primers are given in parenthesis.

Primer	Sequence	Source
SLF-1f	GATCCAGTTGAATGAGATGGA	Tsukamoto et al. 2005
SLF-2f	TGAATGAGATGGAAACTCCTC	Tsukamoto et al. 2005
SLF-3f	AAAGGATGGCGAATGGTA	Tsukamoto et al. 2005
SLF-5f	TGGTTATCCTGGTCCTAAAGATAG	Tsukamoto et al. 2005
SLF-6r	GTGTCAGCGAGTCTTGGAT	Tsukamoto et al. 2005
SLF-7r	GACGAATTTGCACATAAATTTCAGC	Tsukamoto et al. 2005
SLF-8r	GTCCTTGTGATGGTTTGATTG	This study
SGN-f	CCACCAGTATGGATTGTTATTG	This study
SGN-r	AGGTATTTGGGAGAATGATAAG	This study
Sd-f	TTRCCHGAAGATWTKGTGKK (LPED?V?)	This study
Sd-r	YAYTATAAGGTTGTTWGGATT (?YKVV?I)	This study
Sd-fi	AAYCWCTYHTGCGWTTCAAATGTRT (??L?RFKC?)	This study
Sd-ri	AATCCKGCHACMAGARASTTCMGA (NPATR?FR)	This study
SLF-DEGNL1f	TGATGYWYTTAACCCTCTTTTYSC (?DA?NPLF?)	This study
SLF-DEGNL2r	CCDHWAGSKRTTTGGRAGAAW (P?A?W?K)	This study
SLF-DEG3f	CTGGTCCWAAAGAWAGTAAAATTGA (?GPKESKIE)	This study
SLF-DEG6r	CKRTTTGGAAGAASMAYATATTGCT (??WKN?ILL)	This study

Table 3.2. Gen-bank accession numbers for *Petunia* SLF and putative *Solanum SLF* alleles used in this study. The putative *S. tuberosum SLF* sequence was obtained from a BAC Ends sequence on the TIGR Solanaceae Genomics Resource website (www.tigr.org/tdb/sol/sol_ma_blast.shtml).

Allele	Species	Gen-bank ID
PaSLF1	*Petunia axillaris*	AY766155
PaSLF17	*Petunia axillaris*	AY766153
PaSLF19	*Petunia axillaris*	AY766154
PhSLF3	*Petunia hybrida*	AY639402
PhSLFS3A	*Petunia hybrida*	AY639403
PiSLF1	*Petunia inflata*	AY500390
PiSLF2	*Petunia inflata*	AY500391
PiSLF3	*Petunia inflata*	AY500392
SlySLF	*Solanum lycopersicon*	AC171729
StSLF	*Solanum tuberosum*	ER845719

Table 3.3. Primer pairs used in PCR amplifications to isolate the *SLF* gene in *Lycium fremontii*, *L. andersonii* and *Petunia inflata.*

Forward Primer	Reverse Primer
SLF-1f	SLF-8r
SLF-5f	SLF-6r
SLF-5f	SLF-7r
SLF-2f	SLF-8r
SLF-3f	SLF-8r
SLF-2f	SLF-6r
SLF-2f	SLF-7r
SLF-1f	SLF-6r
SLF-3f	SLF-7r
SLF-1f	SLF-7r
SLF-3f	SLF-6r
Sd-f	Sd-r
Sd-fi	Sd-ri
SGN-f	SGN-r
SLF-DEGNL1f	SLF-DEGNL2r
SLF-DEG3f	SLF-DEG6r

Table 3.4. PCR conditions and reagent concentrations for primer combinations. (A) PCR reagent concentrations and amounts are given for different primer pairs for a 25 µl PCR reaction using cDNA. For colony DNA amplifications, 5 µl of DNA is used instead (not indicated on table). Hash marks indicate that several different amounts of a reagent or several PCR programs were used at different times for a particular primer pair. For the SLF-2f and SLF-8r primer pair, amounts with the superscript "N" were used to amplify a putative *SLF* allele after gel extraction. (B) PCR cycles for four PCR programs. For PCR program 4, a touchdown PCR that included annealing steps ranging from 55-50°C was used.

A.

Forward Primer	Reverse Primer	cDNA (ul)	10mM dNTPs (ul)	10uM Primers (ul)	10X Taq buffer	BSA (ul)	25mM MgCl2 (ul)	Taq (ul)	PCR program
SLF-1f	SLF-8r	2.5	0.5	1.25	2.5	0	0	0.2	1
SLF-5f	SLF-6r	2.5	0.5	1.25	2.5	0	0	0.2	1
SLF-5f	SLF-7r	2.5	0.5	1.25	2.5	0.25	0.5	0.25	2
SLF-2f	SLF-8r	2.5/2.5^N/3.5	0.5	1.25/1.25^N/1.88	2.5	0.25	.5/1^N/1.5	0.25	2/3^N/3
SLF-3f	SLF-8r	2.5	0.5	1.25	2.5	0.25	0.5	0.25	2
SLF-2f	SLF-6r	2.5	0.5	1.25	2.5	0.25	0.5	0.25	2
SLF-2f	SLF-7r	2.5	0.5	1.25	2.5	0.25	0.5	0.25	2
SLF-1f	SLF-6r	2.5	0.5	1.25	2.5	0.25	0.5/1	0.25	2
SLF-1f	SLF-7r	2.5	0.5	1.25	2.5	0.25	0.5/1	0.25	3/2
SLF-3f	SLF-7r	2.5	0.5	1.25	2.5	0.25	0.5/1	0.25	2
SLF-3f	SLF-6r	2.5	0.5	1.25	2.5	0.25	0.5	0.25	2
Sd-f	Sd-r	2.5	0.5	1.25	2.5	0.25	0.5	0.25	2
Sd-fi	Sd-ri	2.5	0.5	1.25	2.5	0.25	0.5/*1	0.25	2
SGN-f	SGN-r	2.5	0.5	1.25	2.5	0.25	0.5	0.25	2
SLF-DEGNL1f	SLF-DEGNL2r	2.5	0.5	1.25	2.5	0	0	0.2/.125	2/4
SLF-DEG3f	SLF-DEG6r	2.5	0.5	1.25	2.5	0.25	0.5	0.125	4

B.

PCR Program 1: 1 cycle 94°C (2 min.); 30 cycles 94°C (30 sec.), 45°C (40 sec.), 72°C (1:30 min.); 1 cycle 72°C (10 min.)
PCR Program 2: 1 cycle 94°C (2 min.); 40 cycles 94°C (30 sec.), 45°C (40 sec.), 72°C (1:30 min.); 1 cycle 72°C (10 min.)
PCR Program 3: 1 cycle 94°C (2 min.); 40 cycles 94°C (30 sec.), 40°C (40 sec.), 72°C (1:30 min.); 1 cycle 72°C (10 min.)
PCR Program 4: 1 cycle 95°C (5 min.); 2 cycles 94°C (30 sec.), 55-50°C (1 min.), 72°C (2 min.); 28 cycles 94°C (30 sec.), 49°C (1 min.), 72°C (2 min.); 72°C (10 min.)

Table 3.5. Primer pairs used to amplify the *S-RNase* gene in *Lycium fremontii*, *L. andersonii* and *Petunia inflata*. For different primer combinations, boxes with an "X" indicate that no bands or bands of the incorrect size were obtained from an amplification, boxes with an "N/A" indicate that a primer combination was not used to amplify *SLF* in a particular individual, boxes with "band" indicate that a band of the correct size was obtained for the primer combination used and boxes with "Sequenced" indicate that DNA was cloned and sequenced.

Forward Primer	Reverse Primer	Accessions					
		***Lycium fremontii* 21**	***Lycium andersonii* 9**	***Lycium andersonii* 48**	***Lycium andersonii* 67**	***Lycium andersonii* 80**	***Petunia inflata* 8**
SLF-1f	SLF-8r	Sequenced	X	N/A	N/A	N/A	Sequenced
SLF-5f	SLF-6r	N/A	X	N/A	N/A	N/A	N/A
SLF-5f	SLF-7r	N/A	X	Sequenced	Sequenced	Sequenced	band
SLF-2f	SLF-8r	N/A	X	X	X	X	band
SLF-3f	SLF-8r	N/A	X	X	X	X	band
SLF-2f	SLF-6r	N/A	X	X	X	X	band
SLF-2f	SLF-7r	N/A	X	X	X	X	band
SLF-1f	SLF-6r	N/A	X	X	X	X	band
SLF-3f	SLF-7r	N/A	X	X	X	X	band
SLF-1f	SLF-7r	N/A	X	X	X	X	band
SLF-3f	SLF-6r	N/A	X	X	X	X	X
Sd-f	Sd-r	N/A	X	X	X	X	X
Sd-fi	Sd-ri	N/A	Sequenced	Sequenced	Sequenced	Sequenced	band
SGN-f	SGN-r	N/A	X	X	X	faint	X
SLF-DEGNL1f	SLF-DEGNL2r	N/A	X	X	X	X	X

Table 3.6. NCBI BLAST (Basic Local Alignment Search Tool) results for cloned *Lycium fremontii* and *L. andersonii* DNA sequences. The BLAST results for DNA that included the primer sequence and that had the primer sequence removed are shown. *Lycium* individuals are *L. fremontii* 21, *L. andersonii* 9, 48, 67 and 80. *L. fremontii* DNA was amplified using the primer pair SLF-1f/SLF-8r. DNA from individuals highlighted in gray was amplified using the primer pair Sd-fi/Sd-ri. All other DNA was amplified using the primer pair SLF-5f/SLF-7r. DNA colonies that were cloned from each individual are annotated as "c" and the colony number. Table 3.6 is continued on the next page.

The below table is continued from the previous page.

DNA	BLAST with primer	Gen-bank Accession	BLAST without primer	Gen-bank Accession
L. fremontii 21c2	*S. lycopersicum* DNA sequence from clone LE_HBa-75M11 on chromosome 4, complete sequence.	CU326380	Same	same
L. fremontii 21c8	*Solanum lycopersicum* genomic DNA, chromosome 8, clone:C08HBa0298D02, complete sequence.	AP009291	Same	Same
L. fremontii 21c7	Zebrafish DNA sequence from clone CH211-137C8 in linkage group 4, complete sequence.	BX000521	*Coccidioides immitis* RS hypothetical protein (CIMG_03799) mRNA, complete cds	XM_001244357
L. andersonii 9c4	*Solanum lycopersicum* cDNA, clone: LEFL1089AB05, HTC in leaf	AK247296	Same	Same
L. andersonii 48c7	*Solanum lycopersicum* cDNA, clone: LEFL1089AB05, HTC in leaf	AK24796	*Homo sapiens* chromosome 3 clone CTD-2021F6, complete sequence	AC112216
L. andersonii 67c1	*S.lycopersicum* DNA sequence from clone LE_HBa-75G11 on chromosome 4, complete sequence.	CU459062	*Mus musculus* BAC clone RP24-496E14 from chromosome 13, complete sequence.	AC15239
L. andersonii 80c8	*Solanum lycopersicum* genomic DNA, chromosome 8, clone: C08SLe0082C24, complete sequence.	AP009297	Same	Same
L. andersonii 48c6	*Petunia integrifolia* subsp. *inflata* S2 self-incompatibility ribonuclease (S2-RNase) and S2-locus F-box protein (SLF2) genes, complete cds.	AY136628	*Pan troglodytes* BAC clone CH251-305B15 from chromosome 7, complete sequence.	AC189702
L. andersonii 48c12	*Petunia integrifolia* subsp. *inflata* S2 self-incompatibility ribonuclease (S2-RNase) and S2-locus F-box protein (SLF2) genes, complete cds.	AY136628	*Pan troglodytes* BAC clone CH251-305B15 from chromosome 7, complete sequence.	AC189702
L. andersonii 48c15	*Petunia integrifolia* subsp. *inflata* S2 self-incompatibility ribonuclease (S2-RNase) and S2-locus F-box protein (SLF2) genes, complete cds.	AY136628	*Pan troglodytes* BAC clone CH251-305B15 from chromosome 7, complete sequence.	AC189702
L. andersonii 67c19	*Petunia integrifolia* subsp. *inflata* S2 self-incompatibility ribonuclease (S2-RNase) and S2-locus F-box protein (SLF2) genes, complete cds.	AY136628	*Pan troglodytes* BAC clone CH251-305B15 from chromosome 7, complete sequence.	AC189702
L. andersonii 67c20	*Petunia integrifolia* subsp. *inflata* S2 self-incompatibility ribonuclease (S2-RNase) and S2-locus F-box protein (SLF2) genes, complete cds.	AY136628	*Pan troglodytes* BAC clone CH251-305B15 from chromosome 7, complete sequence.	AC189702
L. andersonii 67c25	*Petunia integrifolia* subsp. *inflata* S2 self-incompatibility ribonuclease (S2-RNase) and S2-locus F-box protein (SLF2) genes, complete cds.	AY136628	*Pan troglodytes* BAC clone CH251-305B15 from chromosome 7, complete sequence.	AC189702
L. andersonii 67c27	*Petunia integrifolia* subsp. *inflata* S2 self-incompatibility ribonuclease (S2-RNase) and S2-locus F-box protein (SLF2) genes, complete cds.	AY136628	*Pan troglodytes* BAC clone CH251-305B15 from chromosome 7, complete sequence.	AC189702
L. andersonii 80c6	*Petunia integrifolia* subsp. *inflata* S2 self-incompatibility ribonuclease (S2-RNase) and S2-locus F-box protein (SLF2) genes, complete cds.	AY136628	*Pan troglodytes* BAC clone CH251-305B15 from chromosome 7, complete sequence.	AC189702
L. andersonii 80c14	Homo sapiens 12 BAC RP11-210N13 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	AC093012	Same	Same

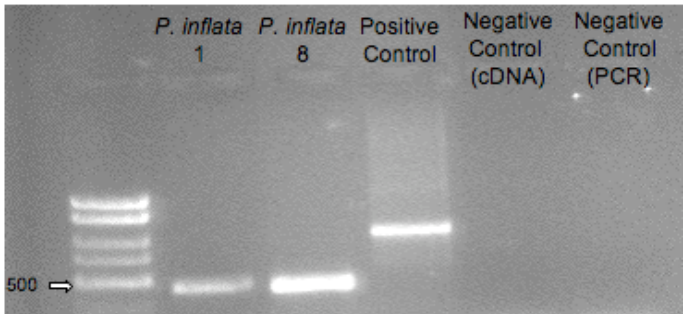

Figure 3.1. Isolation of the *Petunia inflata SLF* gene from cDNA using allele specific primers SLF-1f and SLF-8r. The *P. inflata SLF* gene was isolated from individuals *P. inflata* 1 and *P. inflata* 8. The fifth band down on the DNA ladder is 500 bp.

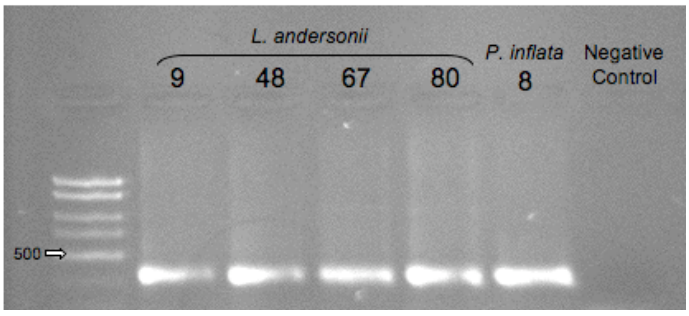

Figure 3.2. Positive amplification of *Lycium andersonii* and *Petunia inflata* cDNA using degenerate primers Sd-fi and Sd-ri. DNA of putative *L. andersonii* and *P. inflata SLF* alleles are shown for *L. andersonii* individuals 9, 48, 87 and 80 and *P. inflata* individual 8. The fifth band down on the DNA ladder is 500 bp.

APPENDIX

RAW DATA ASSOCIATED WITH FIG 1.2

A.1. Raw data associated with Fig. 1.2. Values are average fruit production per flower or average seed number per fruit produced on *Lycium ferocissimum* and *L. pumilum* following outcross or self pollination or in unmanipulated controls. Values in parentheses are standard errors. Sample sizes are given and correspond to either the number of flowers for each treatment (fruit production) or the number of fruits in which seeds were counted (seed number).

	Fruit production			**Seed number**		
Species	**Outcross**	**Self**	**Control**	**Outcross**	**Self**	**Control**
Lycium ferocissimum	0.75 (0.03)	0.08 (0.02)	0.27 (0.03)	43.1 (1.5)	7.9 (2.5)	34.6 (2.3)
	$n = 163$	$n = 174$	$n = 179$	$n = 115$	$n = 14$	$n = 43$
Lycium pumilum	0.60 (0.04)	0.05 (0.02)	0.18 (0.03)	14.7 (0.6)	1.9 (0.7)	4.4 (0.8)
	$n = 136$	$n = 165$	$n = 177$	$n = 71$	$n = 8$	$n = 25$

BIBLIOGRAPHY

Aguilar R, Bernardello G. 2001. The breeding system of *Lycium cestroides*: a Solanaceae with ovarian self-incompatibility. *Sexual Plant Reproduction* 13: 273-277.

Anderson G, Bernardello G, Stuessy TF, Crawford DJ. 2001. Breeding system and pollination of selected plants endemic to Juan Fernández Islands. *American Journal of Bot*ony 88: 220–233.

Baker HG. 1955. Self-incompatibility and establishment after "long-distance" dispersal. *Evolution* 9: 347-348.

Baker HG. 1967. Support for Baker's Law—as a rule. *Evolution* 21: 853-856.

Barrett SCH. 2002. The evolution of plant sexual diversity. *Nat Rev Genet* 3: 275–284.

Barrett SCH, Case AL. 2006. The ecology and evolution of gender strategies in plants: the example of Australian *Wurmbea* (Colchicaceae). *Aust J Botany* 54: 417–433.

Bernardello LM. 1986. Revisión taxonómica de las especies sudamericanas de *Lycium* (Solanaceae). *Boletín de la Academia Nacional de Ciencias Córdoba* 57: 173-356.

Bernardello G, Anderson GJ, Stuessy TF, Crawford DJ. 2001. A survey of floral traits, breeding systems, floral visitors, and pollination systems of the angiosperms of the Juan Fernández islands (Chile). *Bot Rev* 67: 255–308.

Bianchi MB, Gibbs PE, Prado DE, Vesprini JL. 2000. Studies on the breeding systems of understorey species of a Chaco woodland in NE Argentina. *Flora* 195: 339-348.

Bowers JE. 2005. Effects of drought on shrub survival and longevity in the northern Sonoran Desert. *J Torrey Bot Soc* 132: 421–431

Bowers JE, Webb RH, Rondeau RJ. 1995. Longevity, recruitment and mortality of desert plants in Grand Canyon, Arizona, USA. *J Veg Sci* 6: 551–564.

Brennan AC, Harris SA, Tabah DA, Hiscock SJ. 2002. The population genetics of sporophytic self-incompatibility in *Senecio squalidus* L. (Asteraceae) I: S allele diversity in a natural population. *Heredity* 89: 430–438.

Brennan AC, Harris SA, Hiscock SJ. 2003. The population genetics of sporophytic self-incompatibility in *Senecio squalidus* L. (Asteraceae): avoidance of mating constraints imposed by low S-allele number. *Philos T Roy Soc B* 358: 1047–1050.

Brennan AC, Harris SA, Hiscock SJ. 2005. Modes and rates of selfing and associated inbreeding depression in the self-incompatible plant *Senecio squalidus* (Asteraceae): a successful colonizing species in the British Isles. *New Phytol*. 168: 475–486.

Brennan AC, Harris SA, Hiscock SJ. 2006. The population genetics of sporophytic self-incompatibility in *Senecio squalidus* L. (Asteraceae): S allele diversity across the British range. *Evolution* 60: 213–224.

Busch JW. 2005. The evolution of self-compatibility in geographically peripheral populations of *Leavenworthia alabamica* (Brassicaceae). *American Journal of Botany* 92: 1503–1512.

Byers DL, Meagher TR. 1992. Mate availability in small populations of plant species with homomorphic sporophytic self-incompatibility. *Heredity* 68: 353–359.

Carr GD, Powell EA, Kyhos DW. 1986. Self-incompatibility in the Hawaiian Madiinae (Compositae): an exception to Baker's rule. *Evolution* 40: 430–434.

Case AL, Barrett SCH. 2004. Environmental stress and the evolution of dioecy: *Wurmbea dioica* (Colchicaceae) in Western Australia. *Evol Ecol* 18: 145–164.

Castric V, Vekemans X. 2004. Plant self-incompatibility in natural populations: a critical assessment of recent theoretical and empirical advances. *Molecular Ecology* 17: 2873-2889.

Charlesworth D, Vekemans X. 2005. How and when did *Arabidopsis thaliana* become highly self-fertilising. *Bioessays* 27: 472–476.

Chiang-Cabrera F. 1981. A taxonomic study of the North American species of *Lycium* (Solanaceae). Ph.D. thesis, Department of Botany, University of Texas, Austin, USA.

Cruz-Garcia F, Hancock CN, McClure B. 2003. S-RNase complexes and pollen rejection. *Journal of Experimental Biology* 54: 123-130.

de Nettancourt D. 1977. Incompatibility in angiosperms. Springer, Berlin.

Devall MS, Thien LB. 1992. Self-incompatibility in *Ipomoea pes-caprae* (Convolvulaceae). *Am Mid Nat* 128: 22–29.

Entani T, Iwano M, Shiba H, Takayama S, Fukui K, Isogai A. 1999. Centromeric localization of an *S-RNase* gene in *Petunia hybrida* Vilm. Theor Appl Genet 99: 391-397.

Flinn KM. 2006. Reproductive biology of three fern species may contribute to differential colonization success in post–agricultural forests. *Am J Bot* 93: 1289–1294.

Frohman MA, Dush MK, Martin GR. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci USA* 85: 8998-9002.

Goldraij A, Kondo K, Lee CB, Hancokc CN, Sivguru M, Vazquez-Santana S, Kim S, Philips TE, Cruz-Garcia F, McClure B. 2006. Compartmentalization of S-RNase and HT-B degradation in self-incompatible *Nicotiana*. *Nature* 439: 805-810.

Golz JF, Clarke AE, Newbigin E, Anderson M. 1998. A relic S-RNase is expressed in the styles of self-compatible *Nicotiana sylvestris*. *Plant J* 16: 591–599.

Goodwillie C. 1999. Multiple origins of self–compatibility in *Linanthus* section Leptosiphon (Polemoniaceae): Phylogenetic evidence from internal-transcribed-spacer sequence data. *Evolution* 53: 1387–1395.

Hancock CN, Kent L, McClure BA. 2005. The 120k Da glycoprotein is required for S specific pollen rejection in *Nicotiana*. *Plant Journal* 43: 716-723.

Hitchcock CL. 1932. A monographic study of the genus *Lycium* of the western hemisphere. *Annals of the Missouri Botanical Garden* 19: 179-374.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17: 754–755.

Ida K, Norioka S, Yamamoto M, Kumasaka T, Yamashita E, Newbigin E, Clarke AE, Sakiyma F, Sato M. 2001. The 1.55 A resolution structure of Nicotiana alata S(F11) RNase associate with gametophytic self-incompatibility. *Journal of Molecular Biology* 314: 103-112.

Igic B, Kohn JR. 2001. Evolutionary relationships among self-incompatibility RNases. *Proc Natl Acad Sci USA* 98: 13167-13171.

Igic, B, Bohs L, Kohn JR. 2003. Historical inferences from the self-incompatibility locus. *New Phytologist* 161: 97–105.

Igic B, Bohs L, Kohn JR. 2004. Historical inferences from the self-incompatibility locus. *New Phytologist* 161: 97-105.

Igic B, Bohs L, John JR. 2006. Ancient polymorphism reveals univdirectional breeding system shifts. *Proc Natl Acad Sci USA* 103: 1359-1363.

Igic B, Smith WA, Robertson KA, Schaal BA, Kohn JR. 2007. Studies of self incompatibility in wild tomatoes: I. S-allele diversity in Solanum chilense (Dun.) Reiche (Solanaceae). Heredity 99: 553-561.

Ikeda K, Igic B, Ushijima K, Yamane H, Hauck NR, Nakano R, Sassa H, Iezzoni AF, John JR, Tao R. 2004. Primary structural features of the *S* haplotype-specific F-box protein, SFB, in *Prunus*. *Sex Plant Reprod* 16: 235-243.

Ioerger TR, Clark AG, Kao T-h. 1990. Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. *Proc Natl Acad Sci* USA 87: 9732-9735.

Ioerger TR, Gohlke JR, Xu B, Kao T-h. 1991. Primary structural features of the self incompatibility protein in Solanaceae. *Sexual Plant Reproduction* 4: 81-87.

Kao T-h, Tsukamoto T. 2004. The molecular and genetic bases of S-RNase-based self incompatibility. *Plant Cell* 16: S72-S83.

Kipreos ET, Pagano M. 2000. The F-box protein family. *Genome Biology* 1(5): 3002.1-3002.7.

Kondo KM, Yamamoto, Matton DP, Sato T, Hirai M, Norioka S, Hattori T, Kowyama Y. 2002. Cultivated tomato has defects in both *S-RNase* and *HT* genes required for stylar function of self-incompatibility. *Plant Journal* 29: 627–636.

Kumar S, Tamura K, Nei M. 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 5: 150–163.

Lai Z, Ma W, Han B, Liang L, Zhang Y, Hong G, Xue Y. 2002. An F-box gene linked to the self-incompatibility (S) locus of *Antirrhinum* is expressed specifically in pollen and tapetum. *Plant Molecular Biology* 50: 29-42.

Lawrence MJ. 2000. Population Genetics of the Homomorphic Self-incompatibility Polymorphisms in Flowering Plants. *Annals of Botany* 85(Supplement A): 221-226.

Leffler EM. 2006. Ploidy level and self-incompatibility in five *Lycium* species (Solanaceae): evidence from the *S-RNase* locus. B.A. Honors Thesis, Department of Biology, Amherst College, Amherst, Massachusetts, USA.

Levin RA, Miller JS. 2005. Relationships within tribe Lycieae (Solanaceae): paraphyly of *Lycium* and multiple origins of gender dimorphism. *American Journal of Botany* 92: 2044-2053.

Levin RA, Shak JR, Miller JS, Bernardello G, Venter AM. 2007. Evolutionary Relationships in Tribe Lycieae (Solanaceae). *Acta Hort* 745: 225–239.

Lu YQ. 2001. Roles of lineage sorting and phylogenetic relationship in the genetic diversity at the self-incompatibility locus of Solanaceae. *Heredity* 86: 195-205.

Mable BK. 2004. Polyploidy and self-compatibility: is there an association? *New Phytologist162*: 803-801.

Matton DP, Nass N, Clarke A, Newbigin E. 1994. Self incompatibility: how plants avoid illegitimate offspring. *Proc Natl Acad Sci USA* 91: 1992–1997.

Matton DP, Maes O, Laublin G, Xike Q, Bertrand C, Morse D, Cappadocia M. 1997. Hypervariable Domains of Self-Incompatibility RNases Mediate Allele-Specific Pollen Recognition. *Plant Cell* 9: 1757-1766.

Matton DP, Luu DT, Qin X, Laublin G, O'Brien M, Maes O, Morse D, Cappadocia M. 1999. The production of an S-RNase with dual specificity suggests a novel hypothesis for the generation of new S-alleles. *Plant Cell* 11: 2087–2097.

McClure BA, Haring V, Ebert PR, Anderson MA, Simpson RJ, Sakiyama F, Clarke AE. 1989. Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. *Nature* 342: 955-958.

McClure B. 2006. New views of S-RNase-based self-incompatibility. *Current Opinion in Plant Biology* 9: 639-646.

McClure BA, Franklin-Tong V. 2006. Gametophytic self-incompatibility: understanding the cellular mechanisms involved in "self" pollen tube inhibition. *Planta* 224 (2): 233-245.

McMullen CK. 1987. Breeding systems of selected Galapagos Islands angiosperms. *American Journal of Botany* 74: 1694–1705.

McMullen CK. 1990. Reproductive biology of Galapagos Islands angiosperms. *Monogr Syst Bot* 32: 35–45.

Miller JS, Venable DL. 2000. Polyploidy and the evolution of gender dimorphism in plants. *Science* 289: 2335-2338.

Miller JS. 2002. Phylogenetic relationships and the evolution of gender dimorphism in *Lycium* (Solanaceae). *Systematic Botany* 27: 416-428.

Miller JS, Venable DL. 2002. The transition to gender dimorphism on an evolutionary background of self-incompatibility: an example from *Lycium* (Solanaceae). *American Journal of Botany* 89: 1907-1915.

Miller JS, Venable DL. 2003. Floral morphometrics and the evolution of sexual dimorphism in *Lycium* (Solanaceae). *Evolution* 57: 74-86.

Miller JS, Levin RA, Feliciano NM. 2008. A tale of two continents: Baker's law and the maintenance of self-incompatibility in *Lycium* (Solanaceae). *Evolution* in press.

Murfett J, Aherton TL, Mou B, Gasser CS, McClure BA. 1994 S-RNase expressed in transgenic Nicotiana causes S-allele-specific pollen rejection. *Nature* 367: 563-566.

Obbard DJ, Harris SA, Pannell JR. 2006. Sexual systems and population genetic structure in an annual plant: testing the metapopulation model. *American Naturalist* 167: 354–367.

Olmstead RG, Sweere JA, Spangler RE, Bohs L, Palmer JD. 1999. Phylogeny and provisional classification of the Solanaceae based on chloroplast DNA. Pp. 111–137 in M. Nee, D. E. Symon, R. N. Lester, and J. P. Jessop, eds. Solanaceae IV. Richmond Botanic Gardens: Richmond, UK.

Olmstead RG, Bohs L, Abdel Magid H, Santiago-Valentin E, Collier SM, Garcia VF. (in review). A molecular phylogeny of the Solanaceae. Taxon.

Ortiz MA, Talavera S, Garcia–Castaño JL, Tremetsberger K, Stuessy T, Balao F, Casimiro–Soriguer R. 2006. Self-incompatibility and floral parameters in *Hypochaeris* sect. Hypochaeris (Asteraceae). *American Journal of Botany* 93: 234–244.

Paape T, Igic B, Smith S, Olmstead R, Bohs L, Kohn JR. In prep. A 15-Million-Year-Old Genetic Bottleneck.

Pannell JR, Barrett SCH. 1998. Baker's law revisited: reproductive assurance in a metapopulation. *Evolution* 52: 657–688.

Paschke M, Abs C, Schmid B. 2002. Effects of population size and pollen diversity on reproductive success and offspring size in the narrow endemic *Cochlearia bavarica* (Brassicaceae) *American Journal of Botany* 89: 1250–1259.

Posada D, Crandall KA. 1998. ModelTest: Testing the model of DNA substitution. *Bioinformatics* 14(9): 817-818.

Qiao H, Wang H, Zhao L, Zhou J, Huang J, Zhang Y, Xue Y. 2004a. The F-box protein AhSLF-S2 physically ineracts with S-RNases that may be inhibited by the ubiquitin/26S proteasome pathway of protein degradation during compatible pollination in *Antirrhinum*. *Plant Cell* 16: 582-595.

Qiao H, Wang F, Zhao L, Zhou JL, Lai Z, Zhang YS, Robbins TP, Xue YB. 2004b. The F-Box protein AhSLF-S-2 controls the pollen function of S-RNase-based self-incompatibility. *Plant Cell* 16: 2307–2322.

Rambaut, A. 1996. Se-Al: Sequence Alignment Editor. Available at http://evolve.zoo.ox.ac.uk/.

Rambuda TD, Johnson SD. 2004. Breeding systems of invasive alien plants in South Africa: does Baker's rule apply? *Diversity Distrib* 10: 409–416

Raspé O, Kohn JR. 2007. Population structure at the S-locus of *Sorbus aucuparia* L. (Rosaceae: Maloideae). *Molecular Ecololgy* 16: 1315-1325.

Richman AD. 2000. S-allele diversity in *Lycium andersonii*: Implications for the evolution of S-allele age in the Solanaceae. *Annals of Botany* 85: 241-245.

Richman AD, Broothearts, Kohn JR. 1997. Self-incompatibility RNases from three plant families: homology or convergence? *American Journal of Botany* 84: 912-917.

Richman AD, Kao T-h, Schaeffer SW, Uyenoyama MK. 1995. S-allele sequence diversity in natural populations of *Solanum carolinense* (Horsenettle). *Heredity* 75: 405-415.

Richman AD, Kohn JR. 1996b. Learning from rejection: The evolutionary biology of single-locus incompatibility. *Trends in Ecology and Evolution* 11: 497-502.

Richman AD, Kohn JR. 1999. Self-incompatibility alleles from Physalis: Implications for historical inference from balanced genetic polymorphisms. *Proc Natl Acad Sci* USA 96: 168-172.

Richman AD, Kohn JR. 2000. Evolutionary genetics of self-incompatibility in the Solanaceae. *Plant Molecular Biology* 42: 169-179.

Richman AD, Uyenoyama MK, Kohn JR. 1996a. S-allele diversity in a natural population of *Physalis crassifolia* (Solanaceae) (ground cherry) assessed by RT-PCR. *Heredity* 76:497–505.

Richman AD, Uyenoyama MK, Kohn JR. 1996b. Allelic diversity and gene at the self incompatibility locus in the Solanaceae. *Science* 273: 1212-1216.

Rick CM, Fobes JF, Tanksley SD. 1979. Evolution of mating systems in *Lycopersicon hirsutum* as deduced from genetic variation in electrophoretic and morphological characters. *Plant Syst Evol* 132: 279–298.

Rick CM, Chetelat RT. 1991. The breakdown of self-incompatibility in *Lycopersicon hirsutum*. Pp. 253–256 *in* J. G. Hawkes, R. N. Lester, M. Nee, and N. Estrada, eds. Solanaceae III Taxonomy, Chemistry, and Evolution. The Royal Botanic Gardens Press, UK.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

SAS Institute. 1989. SAS/STAT user's guide, version 6, 4th ed. Vol. 2. SAS Institute, Cary, North Carolina, USA.

Savage AE and Miller JS. 2006. Gametophytic self-incompatibility in Lycium parishii (Solanaceae): Allelic diversity, genealogical structure, and patterns of molecular evolution. *Heredity*. 96: 434-444.

Sassa H, Kakui H, Miyamoto M, Suzuki Y, Hanada T, Ushijima K, Kusaba M, Hirano H, Koba T. 2007. *S locus F-box Brothers*: Multiple and Pollen-Specific F-Box Genes With S Haplotype-Specific Polymorphisms in Apple and Japanese Pear. *Genetics* 175: 1869-1881.

Schueller SK. 2004. Self–pollination in island and mainland populations of the introduced hummingbird–pollinated plant, *Nicotiana glauca* (Solanaceae). *American Journal of Botany* 91: 672–681.

Shak JR. 2006. Phylogenetic relationships of Old World Lycium (Solanaceae): Reticulate evolution in the African taxa. B.A. Honors Thesis, Department of Biology, Amherst College, Amherst, Massachusetts, USA.

Sijacic P, Wang X, Skirpan AL, Wang Y, Dowd PE, McCubbin AG, Huang S, Kao T-h. 2004. Identification of the pollen determinant of *S*-RNase-mediated self incompatibility. *Nature* 429:302-305.

Stebbins GL. 1957. Self fertilization and population variability in the higher plants. *American Naturalist* 91: 337–354.

Steinbachs JE, Holsinger KE. 2002. S-RNase-mediated gametophytic self-incompatibility is ancestral in eudicots. *Mol Biol Evol* 19: 825–829.

Stone JL. 2004. Sheltered load associated with S-alleles in *Solanum carolinense.* *Heredity* 92: 335–342.

Stone JL and Pierce SE. 2005. Rapid recent radiation of S-RNase lineages in *Witheringia solanaceae* (Solanaceae). Heredity 94: 547-555.

Sun M, Ritland K. 1998. Mating system of yellow starthistle (*Centaurea solstitialis*, a successful colonizer in North America. *Heredity* 80: 225–232.

Surbanovski N, Tobutt KR, Konstantinovic M, Maksimovic V, Sargent DJ, Stevanovic V, Ortega E, Boskovic I. 2007. Self-incompatibility of *Prunus tenella* and evidence that reproductively isolated species of Prunus have different SFB alleles coupled with an identical *S*-RNase allele. *The Plant Journal* 50: 723-734.

Swofford DL. 2002. PAUP*: phylogenetic analysis using parsimony (* and other methods), version 4b10. Sinauer, Sunderland, Massachusetts, USA.

Tago-Nakazawa M, Dillon MO. 1999. Biogeografía y evolución del clado *Nolana* (Nolaneae-Solanaceae). *Arnaldoa* 6:81-116.

Takebayashi N, Morrell PL. Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. 2001. *American Journal of Botany* 88: 1143–1150.

Takebayashi N. Brewer PB, Newbigin E, Uyenoyama MK. 2003. Patterns of variation within self-incompatibility loci. *Mol Biol Evol* 20: 1778–1794.

Travers S, Mena–Ali J, Stephenson AG. 2004. Plasticity in the self-incompatibility system of *Solanum carolinense*. *Plant Spec Biol* 19: 127–135.

Trouve S, Degen L, Goudet J. 2005. Ecological components and evolution of selfing in the freshwater snail *Galba truncatula*. *Journal of Evolutionary Biology* 18: 358–370.

Tsukamoto T, Ando T, Watanabe H, Marchesi E, Kao T-h. 2005. Duplication of the *S* locus F-box gene is associated with breakdown of pollen function in an S-haplotype identified in a natural population of self-incompatible *Petunia axillaris*. *Plant Molecular Biology* 57: 141-153.

Ushijima K, Sassa H, Dandekar AM, Gradziel TM, Tao R, Hirano H. 2003. Structural and transcriptional analysis of the self-incompatibility locus of almond: identification of a pollen–expressed F–box gene with haplotype-specific polymorphism. *Plant Cell* 15: 771–781.

Uyenoyama MK. 1997. Genealogical structure among alleles regulating self incompatibility in natural populations of flowering plants. *Genetics* 147: 1389-1400.

Vaughan SP, Russell K, Sargent DJ, Tobutt KR. 2006. Isolation of *S*-locus F-box alleles in *Prunus avium* and their application in a novel method to determine self incompatibility genotype. *Theor Appl Genet* 112: 856-866.

Venter AM. 2000. Taxonomy of the genus *Lycium* L. (Solanaceae) in Africa. Ph.D. thesis, University of the Orange Free State, Bloemfontein, South Africa.

Venter AM. 2007. *Lycium hantamense* (Solanaceae), a new species from the Hantam Roggeveld Centre of Plant Endemism, South Africa. *South African Journal of Botany* 73: 214-217.

Wang X, Hughes, AL, Tsukamoto T, Ando T, Kao T. 2001. Evidence that intragenic recombination contributes to allelic diversity of the S-RNase gene at the self incompatibility (S) locus in *Petunia inflata*. *Plant Physiology* 125:1012–1022.

Wang Y, Wang X, McCubbin AG Kao, T-h. 2003. Genetic mapping and molecular characterization of the self-incompatibility (S) locus in *Petunia inflata*. *Plant Molecular Biology* 53: 565-580.

Wang Y, Tsukamoto T, Yi K, Huang S, McCubbin AG, Kao T. 2004. Chromosome walking in the *Petunia inflata* self-incompatibility (*S*-) locus and gene identification in an 881-kb contig containing S_2-RNase. *Plant Molecular Biology* 54:727-742.

Webb CJ, Kelly D. 1993. The reproductive biology of the New Zealand flora. *Trends Ecol Evol* 8: 442–447.

Wilgenbusch, J. 2005. RepMaker 0.9. Available at https://www.csit.fsu.edu/twiki/bin/view/TechHelp/RepMaker/.

Wong WSW, Yang Z, Goldman N, Nielsen R. 2004. Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics* 168: 1041–1051.

Wright S. 1939. The distribution of self-sterility alleles in populations. *Genetics* 24: 538-552.

Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13: 555-556.

Yang Z, Bielawski JP. 2000. Statistical methods for detecting molecular adaptation. *Trends Ecol Evol* 15: 496–503.

Yang Z, Nielsen R, Goldman N, Pedersen AM. 2000. Codon–substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155: 431–449.

Yang ZH, Wong WSW, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol* 22: 1107–1118.

Yeung K, Miller JS, Savage AE, Husband BC, Igic B, Kohn JR. 2005. Association of ploidy and sexual system in *Lycium californicum* (Solanaceae). *Evolution* 59: 2048–2055.